UiT
THE ARCTIC
UNIVERSITY
OF NORWAY

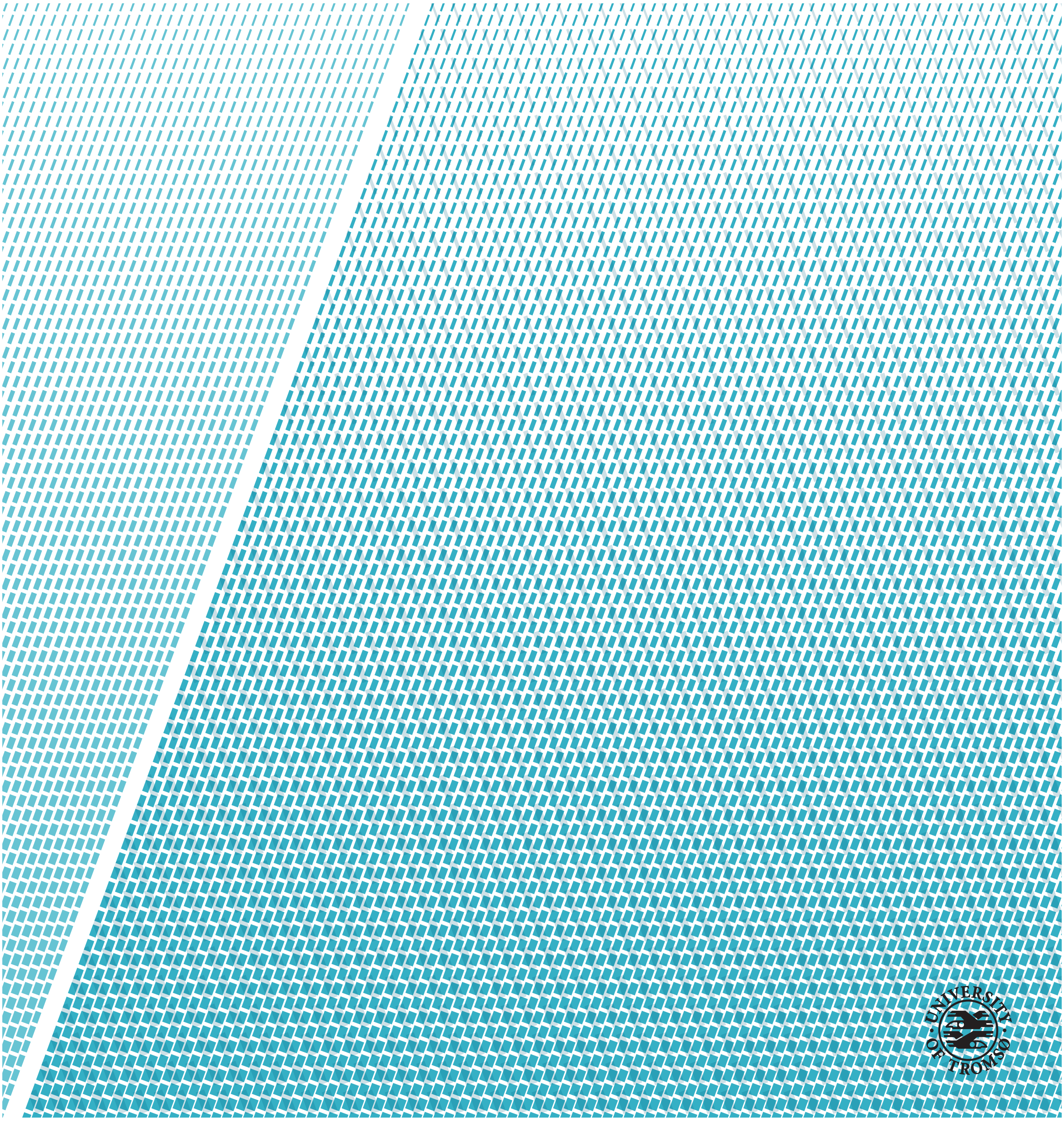
Faculty of Science and Technology
Department of Physics and Technology

Detection and Delineation of Produced Water Slicks in Sentinel-1 Synthetic Aperture Radar Images

---

Brynjar Andersen Saus

EOM-3901 Master's thesis in energy, climate and environment 30 SP June 2021



This thesis document was typeset using the *UiT Thesis LaTeX Template*.

Abstract

Near oil and gas platforms oil detection services regularly detect oil slicks that are a result of legal releases of produced water. These slicks are usually observed using SAR imagery and the important task of observing and monitoring these slicks is as of now carried out manually by human operators aggregated with reported release information. In this thesis we propose three separate approaches to simplify and improve this work through the use of image segmentation and deep learning methods. The approaches are trained and tested on a set of Sentinel-1 scenes over the Brage and Norne platforms off the coast of Norway. The best performing approach was shown to be the direct use of the deep learning algorithm Mask R-CNN on the Sentinel-1 scenes. This approach was able to detect 81% of all slicks in the scenes and had an average user's accuracy of 78% and an average producer's accuracy of 73%. The approaches were also shown to have a significantly reduced ability to detect slicks when the local wind speeds were below 2 m/s or above 11.5 m/s and when the daily volume of oil released from the platforms was below around 150 kg.

Acknowledgements

First, I would like to pay my sincere gratitude to my supervisors, Malin Johansson, Camilla Brekke and Anthony Doulgeris. Your encouragement, guidance and support throughout the last year made this thesis possible.

I also wish to thank my classmates. Striving to keep up with your talent, drive and skill, although at times exhausting, was the main reason I was able to pull through these five years of studies. You have made my time at UiT unforgettable and one of the best experiences of my life. I can't imagine not eating lunch, drinking coffee and annoying you with political monologues every day anymore, but I know we will keep in touch and keep having eachother's backs.

Thank you to my family for always believing in me and supporting me. Mom, dad and Ingeborg, I would not be where I am today without you. You have given me more than anyone could ever ask for, and I am eternally grateful for all of it. I love you.

Finally I would like to thank my friends for always being by my side and supporting me in everything I do. I could not have done it without every single one of you.

The author would like to thank the European Space Agency for providing the Sentinel-1 SAR data. The Sentinel-1 scenes were acquired during the European Space Agency's Copernicus Program and is available here: `https://scihub.copernicus.eu/`.

The author would like to thank Kongsberg Satellites Services for providing reference data and insight into how the reference data is created.

The author would like to thank Norsk Oljevernforening For Operatørselskap for providing optical images of a produced water slick upon request.

The author would like to thank the Norsk Klimaservicesenter for providing weather data.

The author would like to thank Equinor and Winterschall Dea for providing produced water release information for the Brage and Norne oil platforms.

Contents

List of Figures

List of Tables

1

Introduction

In the last decades, Synthetic Aperture Radar (SAR) has emerged as one of the critical components in a multitude of different earth observation and monitoring tasks. SAR has been applied to, among other things, disaster monitoring and management, agriculture and city planning and detection of man-made objects [1–5]. SAR is capable of providing weather and natural illumination independent coverage of the entire earth surface on a weekly basis. This in addition to the unique ability of the radiowaves used to penetrate down into the surface makes SAR useful for a myriad of different applications within the field of earth observation.

SAR has been widely utilized when it comes to marine applications since ground-based observation is limited for ocean areas [6, 7]. In particular, SAR has proved very useful for marine oil spill detection [8]. Oil spills can include accidental releases related to system failure or mechanically damaged equipment, purposeful illegal releases, and legal releases of waste products from oil production. The release of organic compounds and minerals that comes with oil spills of any kind could have significant environmental effects [9, 10]. Early detection and monitoring of mineral oil spills is an integral part of handling and cleanup, and deterring illegal activities.

All films consisting of organic compounds on the ocean surface, including oil slicks, appear as dark patches in SAR images. The low radiation response is a result of the damping effect these types of films have on the capillary and gravitational waves on the ocean surface and the change the film induces in the surfaces dielectric properties [11]. The backscatter from these areas and the ability to differentiate them from the surrounding open water is dependent on weather conditions, oil characteristics, and sensor properties. Figure 1.1 shows a typical example of what an oil slick looks like in a SAR image.

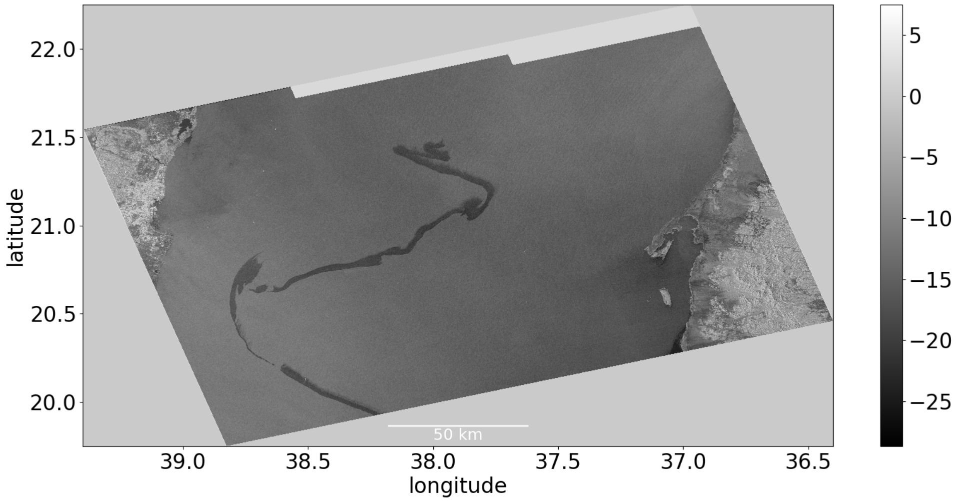


Figure 1.1: This scene shows an oil spill in the Red Sea and is from the Sentinel-1 mission. The image channel is VV and the data has been calibrated to sigma nought, speckle filtered and is shown in decibel. Image: ESA

For decades SAR has been used for monitoring oil slicks through observation by human operators. The current technique of manual detection and delineation demands large amounts of time and resources compared to if these processes were automated. As more SAR data has been made freely available over the years, research on new ways to monitor and detect different types of slicks has significantly increased [12]. In particular, some work has been carried out on automating this process [13–15]. Probabilistic models have been proposed and shown promise for this application [16]. Machine learning approaches have become more viable in recent years with more available data and increased computational power. The deep neural network approach used in [15] was shown to yield results comparable to the ones achieved by human operators.

In SAR images, some features have very similar properties to oil slicks. These features are known as oil slick look-alikes. Look-alikes can be, for example, biogenic films, low wind phenomena, or grease ice. There are also many different oil slicks, including natural leakage, accidental spills from human activity, and purposeful releases of produced water. A recurring issue for automatic detection of oil slicks is differentiating between different films, dark ocean features, and different types of oil slicks. A discussion on the different types of oil slicks and look-alikes and the difficulties of distinguishing them is presented in [17].

Produced water is a byproduct of oil production consisting of water and small amounts of organic compounds and minerals. Oil platforms or other oil-producing installations release produced water into the surrounding ocean as a means of disposal. This activity is legal in most countries but does not come without environmental and ecological concerns [18]. While general oil spill detection in SAR is a longstanding field of study, both for research and operational purposes [19], produced water in SAR is significantly understudied [20–22]. Detection of produced water slicks in SAR shares the difficulties of general oil spill detection, and the slicks are also in general small, short-lived, and due to the low volume and concentration, more sensitive to weather conditions [21]. Wind conditions have a significant effect on the ability to detect produced water slicks as well as other oil slicks. Too low winds make it difficult to separate the slicks from the surrounding waters and too high winds reduced the slicks detectability. On occasion, the slicks are also very narrow, down to only a few pixels across, making delineating and identifying the exact positions challenging. One of the main areas of interest in this thesis will be investigating how these factors affect a machine learning approach to detection and delineation and how to handle these issues.

This thesis proposes automatic, machine learning-based approaches for detecting and delineating produced water slicks in marine SAR images. In addition we propose a semi-automatic approach to simplify the detection work for human operators. The objective is to create flexible approaches, capable of working under different local conditions, that can separate produced water slicks from the surrounding water and differentiate them from look-alikes. The model should be capable of detecting and precisely delineating the slicks while ignoring other features such as natural films and low-wind areas. The work carried out in this thesis will be closely related to earlier work with machine learning-based detection of oil slicks [15, 23]. Building on the approaches described in these papers, it will be attempted to target produced water instead of mineral oil slicks in general. The method will be limited to two classes that divide the images into produced water segments and background. The background will include any feature that is not a produced water slick. This detection of produced water can be combined with semantic detection of other objects in the image to create a complete classification of the whole image, but this is beyond the scope of this work. The main goals of this thesis are:

- Propose automatic or semiautomatic approaches for detecting and delineating produced water slicks in marine SAR images.

- Evaluate environmental and operational conditions that affect the performance of the approaches and the viability of automatic and semiautomatic methods of produced water detection.

- Identify future improvements to data or modelling techniques that could improve automatic detection and delineation of produced water slicks.

The data used in this thesis will consist of SAR images, slick detections and in situ data. The SAR images are from the Sentinel-1 mission and are taken over the Brage and Norne oil platforms off the Norwegian coast between 2018 and 2020. The slick detections are from KSAT and consists of masks of slicks in the images in the dataset. In situ data is provided by platform operators and the Norwegian Meteorological Institute and include local wind speeds and produced water release infomration.

The first part of this thesis introduces the necessary concepts to understand the background for the work carried out. A general description of how SAR works and its most distinctive characteristics is given in chapter 2 and produced water is described and discussed in chapter 3. Chapter 4 introduces the data that will be the basis for the work in this thesis and chapter 5 describes the theoretical background for the proposed methodology. In the second half of the thesis, we go into the specifics of the proposed approaches. The method is described in chapter 6 and the achieved results are shown in chapter 7. In chapter 8 the results are discussed, and potential error sources and shortcomings are presented. Chapter 9 consists of the conclusion to this thesis and 10 outlines some suggestions for future work.

2

Synthetic Aperture Radar

SAR is a a type of *Radio Detection and Ranging* (RADAR) commonly used for earth observation. SAR is an active remote sensing technique that utilizes low-frequency electromagnetic (EM) radiation [24]. This part of the EM-spectrum is known as microwaves or radio waves and consists of radiation with wavelengths of one millimeter to one meter. SAR is an active technique, meaning that it does not depend on naturally present radiation, but instead provides radiation pulses for illumination. This gives SAR the advantage that it is viable independent of the natural lighting conditions of its intended target, making day and night coverage possible. In addition to this, microwaves can penetrate the atmosphere, clouds, and almost any weather condition, bypassing the issues of weather dependency present in optical imagery. Earth observation SAR instruments are commonly spaceborne and satellite-mounted.

2.1 RADAR

RADAR works by transmitting a pulse of EM radiation in the microwave spectrum and measuring the radiation that is reflected back toward the instrument [25]. The pulse can be visualized as an ideal square wave pulse as ilustrated in figure 2.1 a) and b). In this case, and ignoring any attenuation effects, the radiation reflected back toward the instrument will be a convolution of the original pulse. The reflected pulse will also have a time delay compared to the original pulse and the size of this time delay represents the distance from the instrument to the observed object. The attenuation effects are theoretically known to be governed by the radar equation [26]:

$$P_r = \frac{P_t G_t G_r \lambda^2 \sigma}{(4\pi)^3 R^4} \tag{2.1}$$

Where P_r is the received power, P_t is the transmitted power, G_r and G_t are the gain factors for the transmitting and receiving antenna respectively, λ is the wavelength σ is the radar cross section and R is the radius centered on the radar or in other words the distance from the radar to the object. This equation is often rearranged to more precisely denote the different aspects of the signal:

$$P_r = \frac{P_t G_t}{4\pi R^2} \frac{\sigma}{4\pi R^2} \frac{G_r \lambda^2}{4\pi} \tag{2.2}$$

Here each of the three different terms defines a different aspect of the signal. The first term is the effective transmitted power per solid angle and per unit area, the second term is the reflected power per solid angle and unit area in the direction of the receiving antenna and the third term represents the efficiency of the receiving antenna.

Although an ideal square wave can be useful for visualizing a radar pulse, in the real world the pulses are instead approximately triangular pulses as shown in figure 2.1 c) and d).

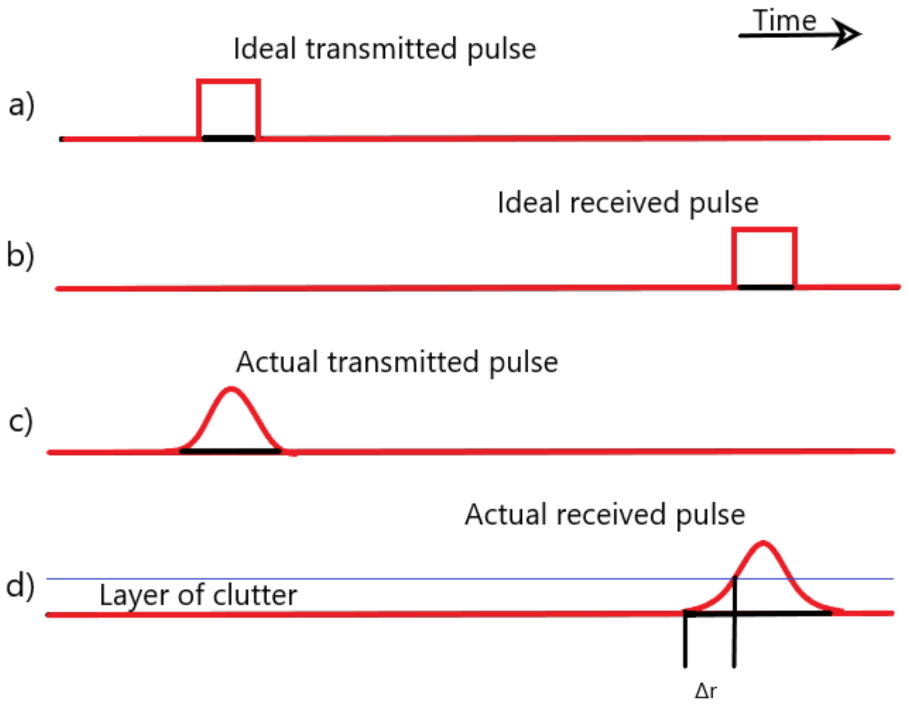


Figure 2.1: Illustration of transmission and reception of radar pulses over time. a) Illustrates a transmitted ideal square wave pulse. b) Illustrates a received ideal square wave pulse. c) Illustrates a transmitted approximate triangular pulse. d) Illustrates a received approximate triangular pulse. Δr is the range inaccuracy and the blue line represents the layer of clutter or unwanted echo registered by the instrument.

The pulse not being an ideal square wave but approximately a triangular pulse has an effect on the spatial resolution. Given that all RADAR systems has a certain amount of noise and clutter in the measured signal the pulse will not be detected before the amplitude surpasses this clutter layer. This results in ambiguity in the exact position in range and two objects with overlapping tails in the received signal will not be discernible for non-coherent radars [25]. The range resolution for a non-coherent radar given by:

$$R_r = \frac{\tau c}{2} \tag{2.3}$$

Where τ is the pulse length and c is the speed of light.

If the EM radiation transmitted is coherent however one can use pulse modulation to improve the range resolution [27]. A phase coherent RADAR is a

RADAR in which the phases of the transmitted pulse is known. The phase of the received signal can then be analysed and compared to the transmitted pulse. Chirp Modulation is the most commonly used pulse modulation. Chirp modulation is also known as linear frequency modulation since a sinusoidal waveform with linearly increasing or decreasing instantaneous frequency over time. Using chirp modulation one can achieve a minimum time resolution, T_{min}, given bandwidth B of:

$$T_{min} = \frac{1}{B} \tag{2.4}$$

Inserting this into equation 2.3 we get a range resolution for a coherent RADAR using a chirp modulation of:

$$R_{r(chirp)} = \frac{c}{2B} \tag{2.5}$$

The visibility of an object for a RADAR is described using what is known as *Radar Cross-Section* (RCS) [28]. RCS is a measure of an object's or target's ability to reflect RADAR signals in the direction of the RADAR Receiver. An object's RCS can be affected by the objects shape, material and angle relative to the RADAR's transmission. RCS, σ, of an object is given by the equation:

$$\sigma = \frac{4\pi r^2 S_r}{S_t} \tag{2.6}$$

Where S_r is the scattered power density in range r and S_t is the power density that is intercepted by the target. When observing an area on the ground that contains multiple objects it is useful to use the related quantity *normalized radar cross-section* or *backscatter coefficient* denoted by σ^0 (sigma nought). This quantity is given by:

$$\sigma^0 = \left\langle \frac{RCS_i}{A_i} \right\rangle \tag{2.7}$$

Where RCS_i is the RCS of object i in the given area and A_i is the area on the ground assisciated with that object.

2.2 Radar with Synthetic Aperture

To create SAR images, consecutive pulses of microwave radiation are transmitted in a side-looking direction from the instrument that is attached to some moving platform, most often a satellite. The backscattered radiation is measured, and the time it takes for the transmitted radiation to return is used to place an illuminated object in the range direction. The longer it takes for some part of the pulse to return, the further away the object reflected off is from the satellite. The sensor needs to be side-looking to avoid any ambiguity in the across-track positioning of objects. Directions in SAR images are defined by the azimuth and range directions. Azimuth signifies the direction parallel to the travel path of the satellite. Range signifies the direction perpendicular to this travel path.

Different surfaces in SAR images can be distinguished because they scatter EM radiation differently [24]. Since SAR images are taken side-looking as illustrated in figure 2.3, what is measured by the instrument is the backscatter from a surface. When EM radiation hits a surface, it can be reflected in two main ways: specular reflection and diffuse reflection as illustrated in figure 2.2. Specular reflection is when the geometric properties of the radiation are preserved. This means that the radiation reflects at an angle from the surface equal to the incidence angle but in the opposite direction. If a surface is a perfectly specular reflector, then all the radiation will be reflected in a single outgoing direction. Diffuse reflection is when the radiation reflects in all directions. A perfectly diffuse reflector is known as a Lambertian surface and will reflect equally in all directions.

Diffuse reflectors that reflect in all directions will have higher backscatter than specular reflectors that reflect more heavily in one direction. In general rough surfaces are more diffuse reflectors while smooth surfaces are more specular reflectors, making rough surfaces appear brighter in SAR images, and smooth surfaces appear darker. A typical example of this relevant to the work in this thesis is the relationship between clean ocean water and produced water slicks. Produced water creates a smoother surface which therefore appears darker than the rougher surrounding open waters.

Figure 2.2: Illustration of specular and diffuse reflection.

SAR differentiates itself from the earlier method, Real Aperture Radar (RAR), by using a synthetic aperture [24]. RAR only employs the antenna statically and merely sends out a pulse of radiation and measures what is scattered back. With RAR, the spatial resolution in the azimuth direction of the image produced is limited by the size of the antenna. For practical reasons, the use of a giant antenna to get improved resolution is not viable. SAR circumvents this by simulating a larger antenna using the movement of the platform. This makes it possible to get images with higher spatial resolution without needing to employ massive antennas. The azimuth resolution of RAR is given by:

$$R_a = \frac{\rho\lambda}{L} \tag{2.8}$$

Where ρ is the nominal slant range, λ is the wavelength and R_a is the azimuth resolution. We can see from this that the resolution improves with the a larger antenna. Extrapolating on this equation we can find the azimuth resolution in SAR. Given that the nominal slant range can be written as $\rho = \frac{H}{\cos\theta}$ we get an azimuth resolution given as:

$$R_a = \frac{\rho\lambda}{L} = \frac{\lambda H}{L\cos\theta} \tag{2.9}$$

Now by forming an antenna with length equal to $2R_a$ we get an azimuth range for SAR:

$$R_a = \frac{\lambda\rho}{2R_a} = \frac{L}{2} \tag{2.10}$$

As we can see from equation 2.10 the azimuth resolution gets better with reduced antenna size for SAR. So by the employment of this principle one circumvents the limitations of large antennas i RAR. The lower limitation on antenna size is the increased noise resulting from a smaller antenna [24].

Range resolution of SAR is determined by the frequency bandwidth of the transmitted pulse. This implies that it is affected by the time duration of the range focused pulse. Large bandwidths yields a small pulse width and vice versa.

When working with SAR one differentiates between slant range and ground range [29]. Slant range is the natural result produced by the measurements made by the RADAR instruments. Ground range is the result of correcting for local terrain slope and elevation. Slant range only informs on the location of objects in the image relative to the instrument, while ground range places the objects in the image relative to eachother. Figure 2.3 illustrates ground range and slant range. The slant range resolution is given in equation 2.3. The ground range resolution is on the other hand is also dependent on the look angle and is given by:

$$R_r = \frac{C\tau}{2}\frac{1}{\sin\theta} \tag{2.11}$$

Where R_r is the ground range resolution, and θ is the incidence angle. The slant range is the resolution given the observed surface being perpendicular on the incident radiation beam and the ground range is the actually observed resolution when the surface is at an angle from the beam.

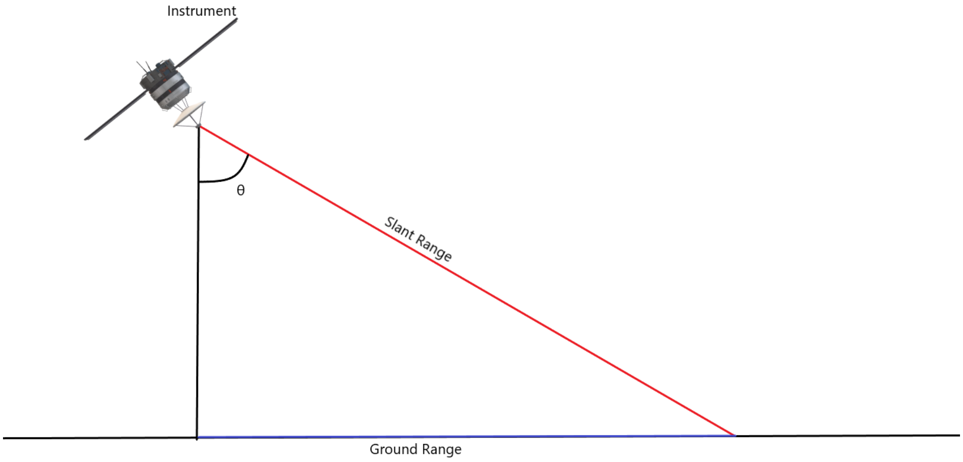


Figure 2.3: Illustration of slant range and ground range. θ is the incidence angle.

The microwaves and radio waves that SAR employs are divided into bands depending on frequencies. Different bands have different properties and different applications. The wavelengths mainly applied in SAR imagery are the X-, C-, S-, L- and P- bands.

Table 2.1: The different microwave-bands used in SAR.

Band	Wavelength (cm)	Frequency (GHz)	Typical Applications
X	2.4-3.8	8-12	Urban monitoring, vegetation, snow and ice
C	3.8-7.5	4-8	Global mapping, Ocean maritime navigation
S	7.5-15	2-4	Agriculture monitoring
L	15-30	1-2	Medium resolution SAR
P	30-100	0.3-1	Biomass, Vegetation mapping and assessment

2.3 Modes of SAR

SAR satellites can generally operate in three different acquisition modes. These are the Stripmap (SM) mode, the scanning SAR (ScanSAR) mode, and the spotlight mode. Each mode has its advantages, disadvantages, and applications [30].

In SM acquisition mode, the instrument is stationary and looks in a fixed direction from the platform it is attached to. As the instrument moves, it obtains continual imaging of the area the satellite moves over. SM provides continual coverage over large areas. It creates vast amounts of data, so it is only used in special situations where continual coverage over a large area is necessary, such as in emergency management [30].

In the ScanSAR acquisition mode, the instrument sweeps over the target area to create subswaths in the range direction. These subswaths are then combined to create the whole image. ScanSAR makes it possible to increase the swath width covering a larger area than other modes such as SM. However, this makes the data more computationally complex and can result in some ambiguities around the merging of the subswaths. This is particularly noticeable when using polarizations with generally low amplitudes such as VH over maritime areas. ScanSAR can also suffer from some edge effects between two beams when they are spliced together and scalloping [31] as seen in figure 2.4. These effects can complicate analysis and object detection.

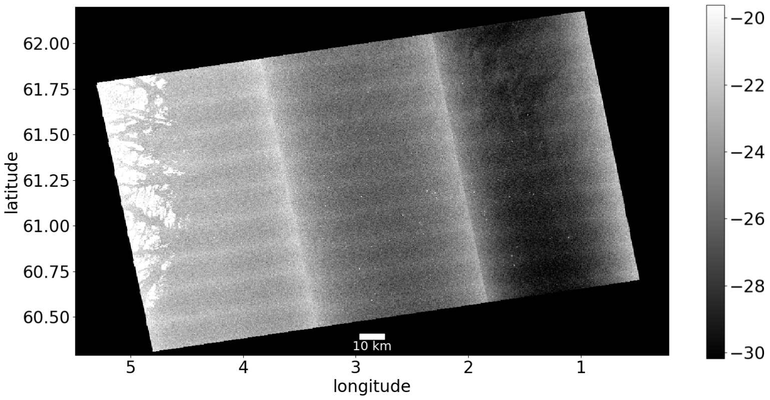


Figure 2.4: VH-channel Sentinel-1 scene over the Brage oil platform with clear edge effects along the connection of the different beams and scalloping effects. The image is presented in decibel.

When needing higher resolution images of one specific location, one would use the spotlight mode. In spotlight mode, the beam is controlled to focus on one spot for a longer duration of time while passing it. By keeping the beam fixed on one area, one can create a longer synthetic aperture. While this does improve the resolution of the image, it comes at the cost of spatial coverage.

2.4 Polarization

An important aspect of SAR data is what is known as polarization [32]. Polarization refers to the geometric orientation of the oscillations in the electric field of EM radiation. Vertical polarization means that the field oscillates up and down when seen from straight ahead, while horizontal polarization means that the field oscillates left to right, see figure 2.5. The polarization of EM radiation is divided into linear polarization, circular polarization, and elliptical polarization. For this thesis, only linear polarization is relevant. Linear polarization is differentiated from the other two types by having a constant electric and magnetic field direction. In contrast, circular and elliptical polarization have a constant rotation of the fields.

Polarization allows us to gather more information from SAR imagery. By controlling the polarization of the pulse from the instrument and the polarizations that it can detect, one can observe the surface’s effect on polarization. The backscattered radiation may have the same polarization as the original pulse or the opposite. This provides us with four potential channels known as HH, HV,

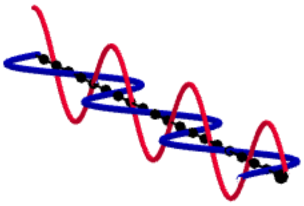


Figure 2.5: Polarization of electromagnetic radiation. The red line is vertically polarized and the blue is horizontally polarized.

VH, and VV. H for horizontal and V for vertical, where the first letter denotes the transmitted polarization and the second denotes the received polarization. The different channels will have varying responses to different surfaces providing the observer with additional information compared to using only one channel.

When using polarization in this way, it is known as polarimetric SAR. Polarimetric SAR can be divided into three types based on how many different channels are available. Single-polarization (single-pol) means that the instrument transmits and receives only one polarization. Dual-polarization (dual-pol) means that the instrument only transmits in one polarization but can receive both types. Quad-polarization (quad-pol) is the most complete form of polarimetric SAR. In quad-pol, the instrument can transmit and receive both polarizations resulting in all four channels being available.

Using multiple polarizations does provide multi-dimensionality, which could increases the precision of classification as shown in [33]. This is to be expected, as increasing the amount of data used for classification should improve the separability between classes [34]. However, using polarimetric data comes with some challenges as well [35]. The increased amount of data makes analyses more computationally demanding. This can make it necessary to view smaller images by either reducing resolution or area covered. It is also worth noting that the magnitude of the different polarizations are not equal. Usually, the response in the cross-polarization is low compared to in the co-polarization, especially for marine images. This might make the magnitude lower than the noise, causing this part of the data to contribute little of interest to the dataset. The dataset used in this thesis consists of dual-pol SAR images. The two channels available are VV and VH.

2.5 Incidence Angle Degradation

The pulse radiation from the SAR instrument will strike the ground at different angles depending on where in the range direction we are looking. This is illustrated in Figure 2.6. From Lambert's cosine law [36] we know that the observed radiant intensity of a surface is directly proportional to the cosine of the angle between the direction of the incident radiation and the normal of the surface.

$$I_d = k_d I_{incident} \cos\theta \tag{2.12}$$

Where I_d is the observed radiant intensity, k_d is the diffuse reflectance of the surface, $I_{incident}$ is the incident radiation on the surface and θ is the incidence angle.

This implies that the observed backscattered radiation will be higher in the near-range part of the image where the incidence angle is steeper and lower in the far range. This effect can emulate differences in surface types and create erroneous classifications when segmenting or classifying different areas in an image.

It has also recently been shown that the incidence angle degradation effect is not equal for all types of surfaces [37]. Different surfaces will have different degrees of degradation, and the impact of the effect may vary significantly between them. This makes the degradation more challenging to map and account for when processing and analyzing SAR images. However, the effect is consistently approximately exponential in power, meaning that it is linear in decibels with different slopes for different surfaces. We can see the degradation in power and decibels in the graphs in figure 2.6. Figure 2.7 shows an example of different incidence angle degradations for different surface types.

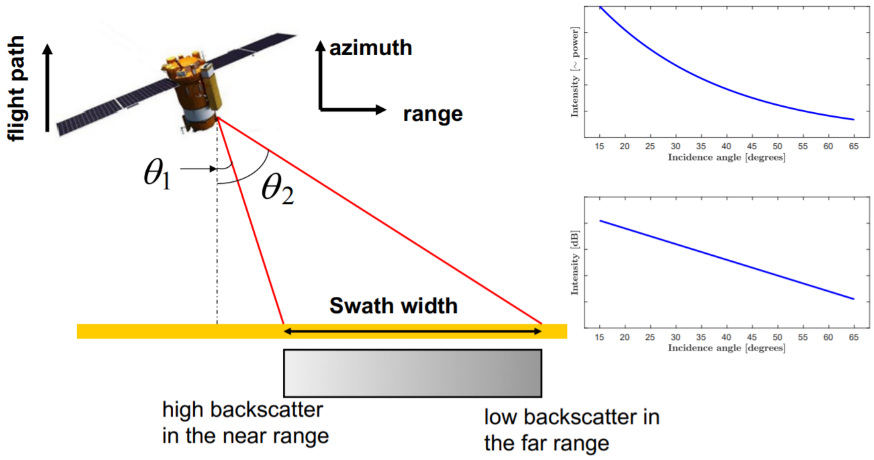


Figure 2.6: This is an illustration of how SAR satellites work. SAR sensors are side-looking, affecting the incidence angles for different parts of the swath. The intensity reduction is approximately exponential in power or linear in decibels, as seen in the graphs in the figure. Used with the permission of Anthony Doulgeris [38].

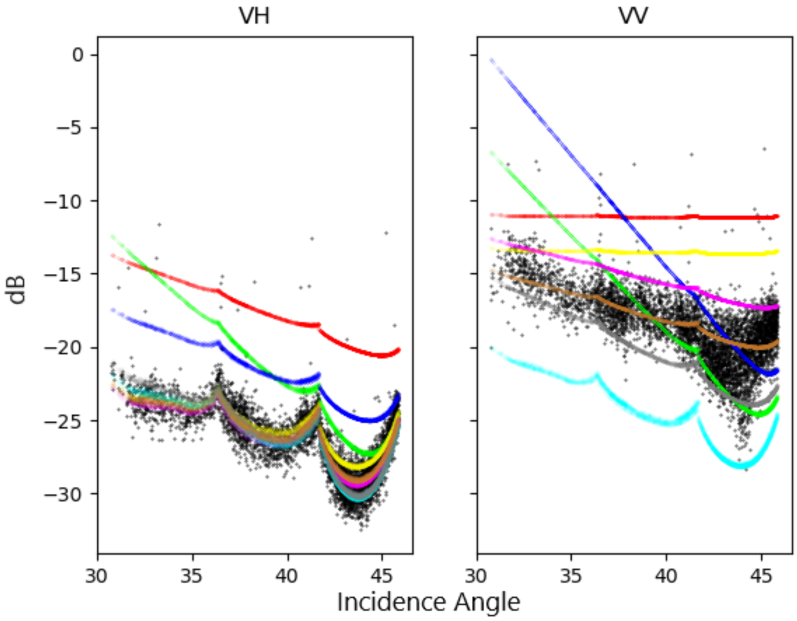


Figure 2.7: Plot of incidence angle degradations for different surface types in a Sentinel-1 marine SAR image. Each line represents a different class of the image as determined by the segmentation algorithm described in section 5.2. The x-axis consists of the incidence angle and the y-axis is the intensity of the noise floor.

2.6 Noise

Noise in a signal is any unwanted interference that is collected in addition to the signal itself. The main types of noise in SAR images are speckle noise and thermal noise.

Thermal noise refers to the noise generated by thermal agitation of the sensor itself when collecting data [39]. When the charge carriers in the sensor are thermally agitated, this is registered as a signal input in the sensor and becomes part of the produced data. Thermal noise is existent in any electrical component but is more prominent in higher temperatures. The power of thermal noise in a sensor is given by the equation:

$$N = k_b T \delta f \tag{2.13}$$

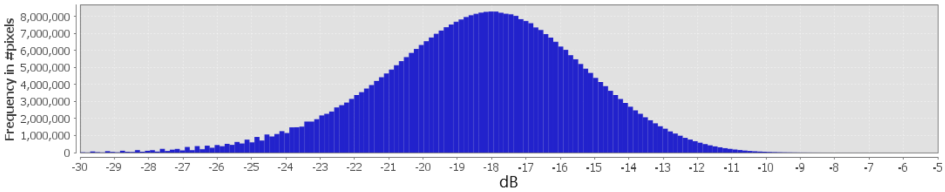
Where N is the noise power, k_b is the Boltzmann constant, T is the temperature, and δf is the bandwidth.

The most dominant form of noise in SAR imagery is speckle noise. Speckle is a type of interference in the signal characterized by the image being speckled with very dark and very bright pixels. This interference results from the coherent nature of SAR imaging and the random position of individual scatterers within each resolution cell. The produced signal is complex, consisting of both magnitude and phase, and it can interfere both constructively and destructively. This is why the pattern created consists of both very dark and very bright spots. Even though speckle is an unwanted part of the signal that needs to be accounted for when analyzing the images, it is not noise in the traditional sense. It is instead an inherent part of the measured signal.

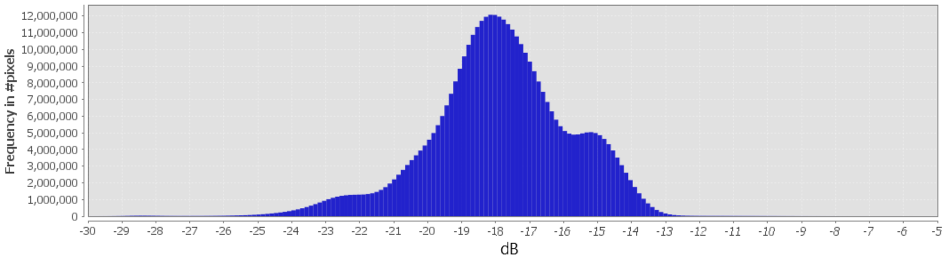
Figure 2.8 shows a good example of how speckle noise makes an image more challenging to interpret statistically. In figure 2.8a the signal looks like it can be best represented as one Gaussian distribution. This implies that a Gaussian mixture model segmentation algorithm, such as the one that will be used in this thesis, would group all the pixels in the same segment [37]. In figure 2.8b the speckle noise has been removed, and we can see that the signal consists of at least three distinct classes.

An important parameter for evaluating the quality of a signal is the *signal-to-noise ratio* (SNR). SNR is a measure of the relationship between the noise levels in a signal and the strength of the wanted signal. SNR is represented in decibel and is given by:

$$SNR = 10 \log_{10}(\frac{P_s}{P_n}) \tag{2.14}$$



(a) Before speckle filtering



(b) After speckle filtering

Figure 2.8: Histograms of the VV intensity of a SAR image before and after speckle filtering. The image has been σ_0 calibrated in both cases, and intensity values are in decibel.

Where P_s is the power of the signal and P_n is the power of the noise. If the SNR is too low either because of the signal having low power or the noise having high power, the signal becomes more difficult to interpret. Another parameter used to measure the sensitivity of a system is the *Noise-Equivalent Sigma Zero* (NESZ). NESZ is a measure of how sensitive the system is to area of low backscatter and is given by the value of the backscatter coefficient corresponding to a SNR equal to one.

3

Produced Water

Produced water is a waste product from oil and natural gas extraction. In the process of extraction there are two main sources of produced water [40]. The first is water that is already present in the subterranean pockets containing the oil and natural gas mixing with the product as it is being extracted. The second is the water pumped into the reservoir that is ultimately retrieved as produced water. Produced water is a brackish and saline water and pollutants mixture containing around 300-900 g/m^3 organic carbons before cleaning [41]. In addition to containing organic compounds from the oil itself, produced water often contains metals, production treatment chemicals and other pollutants. The exact composition of the produced water depends on the technologies used, the age and depth of the geological formation and the geographic location.

3.1 Treatment and Disposal

Due to environmental concerns most countries have strict regulations detailing how produced water needs to be handled [42]. This has made development of treatment techniques a valuable investment and many different technologies have been developed over the years [43]. So far no viable treatment that can remove all hydrocarbons and other impurities from the produced water has been created. This results in the need for disposal of the remaining waste products. There are many possible approaches to this task [44] and the regulating bodies of each oil producing country set the standards for the disposal process.

The two platforms that will be investigated in this thesis are the Brage platform and the Norne platform. They are both under Norwegian jurisdiction and therefore the most relevant subject for this work are the Norwegian government's rules and regulations for the handling of produced water. In Norway the law prohibits release of produced water with a monthly mean oil content over 30 g/m^3 [45] . Any party that is going to release produced water in Norwegian waters are required to perform an environmental evaluation regarding the effects and consequences of the release. It is also required that this evaluation is repeated if the volume or concentration changes or every five years barring any changes.

3.2 Produced Water in SAR

Water is denser than oil and oil and water molecules do not coalesce. This results in the oil creating a film on top of the water surface. Since the oil molecules are attracted to the water molecules this film spreads out and becomes relatively thin, covering a large area. One of the main reasons SAR is so often applied in oil slick detection and observation is the very distinct and recognizable look of these features in radar images [46]. Oil slicks appear as dark patches in SAR images. This is because the backscatter from the surfaces covered by an oil slick are lower than that of the surrounding waters. This is the case for accidental spills, natural leakage and produced water. There are two main drivers for the intensity contrast between slicks and open water; the dampening effect that the film has on the ocean surface and the change in dielectric properties in oil slicks compared to clean water [11].

Radar backscatter caused by surface roughness is mainly a result of Bragg resonance scattering [47]. Wind causes spatially-correlated ocean wave components, of which those that are in phase with the incident radiation contribute to resonant scattering. In addition, the effects of gravitational pull from celestial bodies and gravitational differences distributed around the earth contribute to causing waves on the ocean surface [48]. The dampening effect is a result of the film counteracting the capillary and gravitational waves of the surface, removing its inherent roughness. Oil released into the water will settle on the surface and create a film. This film causes dissipation of the wave energy by Gibbs surface elasticity and a reduction in the surface tension [49]. The energy contained in the waves is partially absorbed by the monolayer surface elasticity and the reduced surface tension reduces the energy transfer between adjacent surfaces. The film also reduces wind drag on the surface, resulting in less energy transfer from the wind to the water surface. This in turn combined with the other wave-related effects of the film leads to less waves and a smoother surface. This effect is illustrated in figure 3.1, where we can see how a film forms on the water surface and contributes to reducing surface waves.

The film also affects the dielectric properties of the surface. This change in dielectric properties have an impact on the interaction between the surface and the electromagnetic waves from the SAR-instrument. Oil has a lower relative dielectric constant than seawater. The relative dielectric constant is around $\epsilon_r^O = 2.3 - i0.02$ [50] for oil and around $\epsilon_r^{SW} = 76.01 - i51.71$ for seawater in typical North Sea conditions (T=5°C, S=34.992) [51]. The emulsion between seawater and the oil slick will have a relative dielectric constant that lies between ϵ_r^O and ϵ_r^{SW}, in other words, one that is lower than that of the clean seawater. A reduced dielectric constant means that the total energy that the surface reflects is decreased [11]. Since the total reflected energy from the surface will be lower for the slick compared to the surrounding waters, the backscatter measured by the SAR instrument will also be lower, making the slick appear darker in SAR images.

For the change in relative dielectric constant to have an effect on the measured backscatter, the film must be thick [50, 52]. Figure 3.1 depicts how a slick affects the dielectric properties of the water surface. If the oil slick forms only a very thin layer or a sheen on the surface, the radiation passes through the oil and reflects of the seawater underneath. In such a case the effective relative dielectric constant for the surface will be the same at the relative dielectric constant of clean water. Produced water, due to its low oil concentration, forms very thin films on the surface [21]. It is therefore likely that the effect of changes in dielectric properties are small and the dominating contrast driver for produced water slicks is the dampening effect of the slicks on the capillary and gravitational waves. This is also supported by the fact that wind conditions have a large effect on detectability [21] and the damping ratio being close to constant across each individual slick. Figure 3.1 shows that for a thicker slick both wind dampening and dielectric effect the backscatter, while for a sheen as is most often the case for produced water, only the wind dampening has an impact.

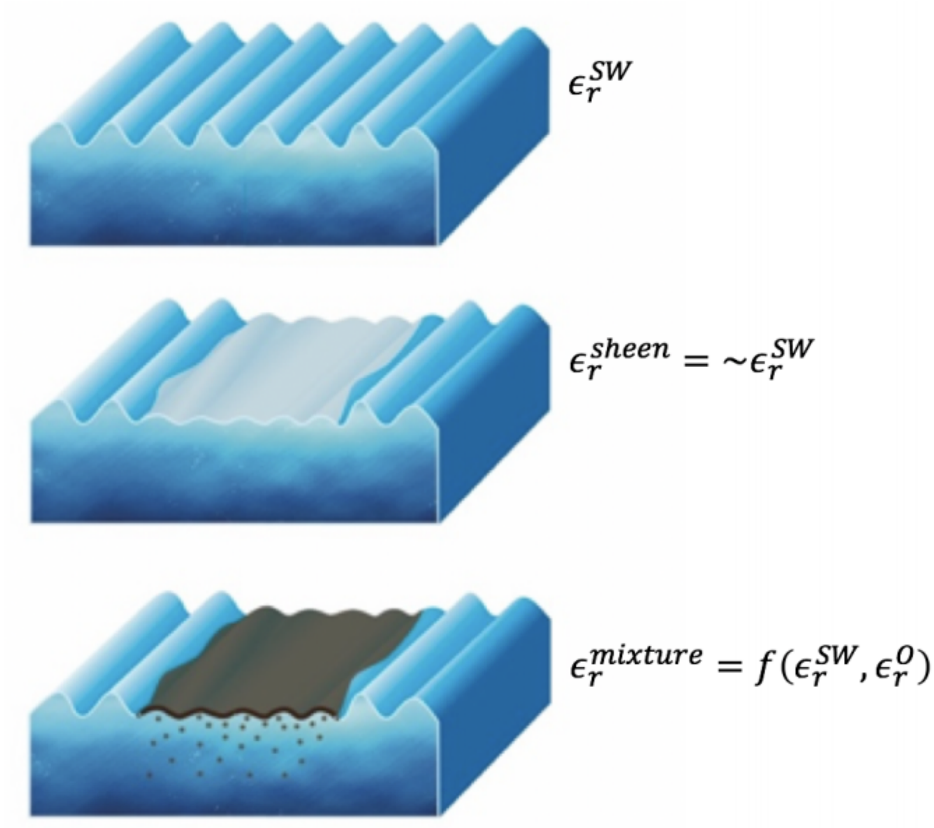


Figure 3.1: The effect of a sheen and a thicker oil slick on a seawater surface. The first illustration shows clean ocean water, the second shows a thin sheen and the third shows a thicker slick. Figure used with the permission of Camilla Brekke [11]

The Bonn Agreement [53] has defined a color spectrum for classifying the thickness of a slick based on the color and look in optical imagery. This has been used by researchers to get information on the thickness of a slick based only on earth observation approaches, without needing to rely on on-site measurements [54]. A slick classified as a sheen, that will be too thin for the dielectric properties of the surface to change, is defined in the Bonn agreement as looking grey or silvery. Norsk Oljevernforening For Operatørselskap (NOFO) has for the purposes of this research provided video over the Brage platform containing a produced water slick. A screenshot from this video can be seen in figure 3.2. In the image we can see the slick as the grey area stretching from top to bottom in the middle. In this image we can clearly see that the slick looks silver colored. This suggests that the slick is a sheen and therefore can conclude that the dielectric properties will remain nearly unchanged.

Figure 3.2: Screenshot from video provided by NOFO doing a flyover over the Brage platform. The grey area in the middle of the image that stretches across the image from top to bottom is an example of a produced water slick.

As a result of these contrast drivers slicks are visible in SAR images as dark patches contrasted with the usually much brighter surrounding waters. How large the contrast is has a significant effect on the visibility of the slicks. The relation between the backscattered power from oil slicks and the surrounding waters is known as the *damping ratio* [21]. This is a measure of how much the backscatter of the surface is dampened by an oil slick. The damping ratio can vary both with the roughness of the surrounding waters and with the smoothness of the oil slicks. It has been shown that the damping ratio increases both with increased winds up to 10 m/s and with increased thickness of the film [55]. It has also been shown in literature that the damping ratio is useful in differentiating different type of oil spills, determining whether an oil spill will be detectable using SAR alone and other applications such as investigating temporal changes [56, 57]. The damping ratio, ζ, is given by:

$$\zeta = \frac{\sigma_{\text{cleansea}}}{\sigma_{\text{slick}}} \tag{3.1}$$

Where σ_{cleansea} is the average intensity of an area with open water and σ_{slick} is the average intensity of the slick or a part of the slick.

An increase in damping ratio would indicate that the slick is more separable from the surrounding waters. The damping ratio is not a constant measure for any one type of slick or even any one scene. As different areas in a scene could be subject to different environmental conditions, the damping ratio could vary. The damping ratio can in such a case be measured against an average value for the intensity values of the rest of the scene or some chosen area of interest. The most relevant is often to evaluate the damping ratio based on the slick and the open waters adjacent to it. This will give the clearest indication of how easily a slick is detected in the image. The slicks in this dataset have damping ratios of around 3–8.5dB. This is consistent with what has been found in earlier works [21].

Although produced water slicks share the low backscatter property with all oil slicks and oil slick look-alikes, there are some unique characteristics for produced water slicks that make them possible to differentiate. Being the result of controlled disposal from an oil platform, produced water slicks always originate from one stationary point. This point is the platform, which is also often clearly visible in SAR imagery because of the high backscatter from man-made structures [58]. In addition the low concentration of hydrocarbons in the produced water and the often relatively low volumes that are released result in a slick that is much smaller than for example spillage from a large oil tanker, making it distinguishable from these larger mineral oil spills [21]. In figure 3.3 we see a typical produced water slick from the Brage platform. Here we can see the dark slick originating from the bright spot that is the platform.

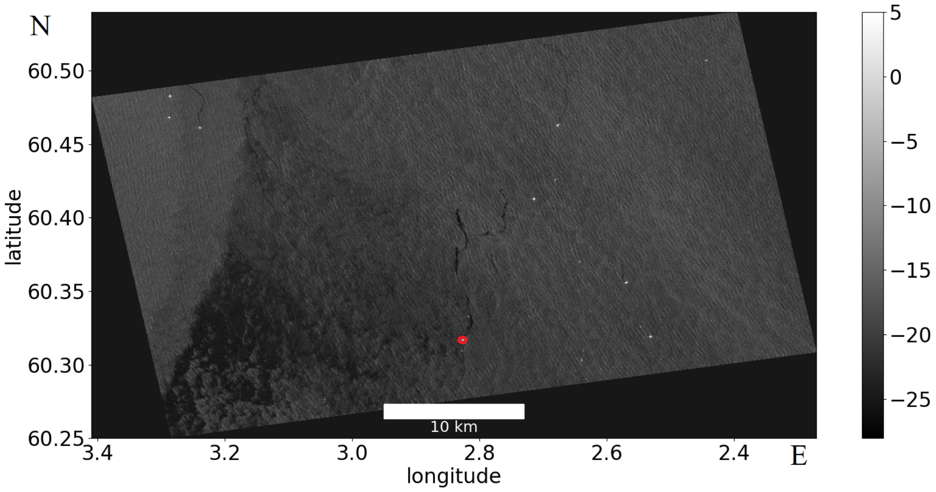


Figure 3.3: Scene with a produced water slick from the Brage platform. The image is in the VV-channel and in decibels. The slick begins in the middle of the bottom of the image. Brage is marked with a red circle

4

Data

This chapter contains an introduction to, and a general description of the dataset used in this thesis's work. In chapter 8 one can find a more in-depth discussion on how the characteristics of this data might affect the results.

Like many other earth observation research projects, this thesis is made possible by the increases in freely available satellite data over the last decades. The introduction of an open and accessible policy from a series of satellite owners has heralded an explosion in scientific inquiry on the subject [12]. Freely available SAR data from satellites has contributed to revolutionizing earth observation in general and particularly machine learning approaches since large quantities of data are now attainable for scientists worldwide.

4.1 Study Site

Norway has since the 1980's been a big producer of fossil fuels due to the large oil and gas reservoirs off the coast of the country. The Norwegian continental shelf has over the decades been speckled with a significant number of drilling sites and oil platforms. All these platforms need to dispose of produced water, making produced water slicks very prominent along almost the entire Norwegian coast. We can see the installations off the coast of Norway in figure 4.1

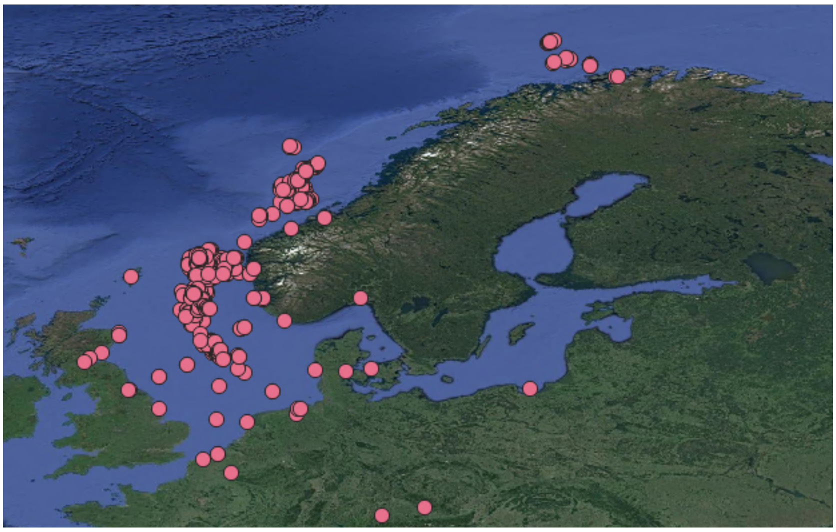


Figure 4.1: Locations of all fixed facilities and floating production facilities and main facilities onshore under the Norwegian Petroleum Directorate.

The dataset used for this thesis consists of scenes containing produced water slicks from two oil platforms, Brage and Norne near the coast of Norway. Figure 4.2 shows the location of both platforms. Norne is a floating platform, and Brage a stationary platform. Both platforms continually release produced water into the surrounding ocean and the slicks created by these releases are often visible in SAR imagery and regularly detected by oil spill detection services.

The satellite scenes are chosen specifically to include the Brage and Norne platforms, but because of the frequency of oil and gas producing installations along the Norwegian coast, other platforms also fall within the satellite scenes. These platforms also dispose of produced water and therefore also produce slicks in the images, resulting in some scenes having more than one instance of detected oil slicks. This is however not necessarily a weakness, as more instances of slicks in the scenes result in more varied training data and a better test for the flexibility of the approaches proposed in this thesis.

Figure 4.2: Norne and Brage locations off the coast of Norway. Brage is located at 60.5425◦N, 3.0468◦E and Norne is located at 66.0138◦N, 8.0158◦E

4.2 SAR Images

Within this thesis SAR images from the Sentinel-1 mission is used. Sentinel-1 is a satellite constellation consisting of two satellites with two more planned belonging to ESA [30]. The data from this mission is freely available to anyone. The current satellites in the constellation are the Sentinel-1A launched in 2014, and Sentinel-1B launched in 2016. The two satellites are in a polar orbit 180°apart. The constellation orbits at an altitude of 700km and has a global revisit time of 6 days [30]. Sentinel-1 gathers data using a C-band SAR instrument as described in table 2.1. The dataset used in this thesis consists of 388 images taken in 2018, 2019, and 2020 that cover the operational Brage and the Norne oil platforms. Both platforms are located off the coast of Norway with coordinates 60.5425◦N, 3.0468◦E for Brage and 66.0138◦N, 8.0158◦E for Norne.

Sentinel-1 employs four different acquisition modes; SM, Interferometric Wide Swath (IW), Extra Wide Swath (EW), and Wavemode (WV). Sentinel-1 uses an improved version of ScanSAR known as Terrain Observation with Progressive

Scans SAR (TOPSAR). In TOPSAR, the beam can be steered backward and forward in the azimuth direction and the range direction. This contributes to avoiding scalloping and creates a uniform image quality throughout the swath. Sentinel-1 has two different operating modes that use the TOPSAR technique, IW and EW. IW is the primary acquisition mode over land and coastal areas, while EW is mainly used to observe sea-ice, over polar areas, and a few other maritime applications. The general operating mode used by Sentinel-1 over open ocean is WV. In this mode, the beam skips in a "leapfrog" acquisition pattern. The data is collected as non-adjacent images with ample spacing between them.

An overview of the different acquisition modes can be seen in figure 4.3.

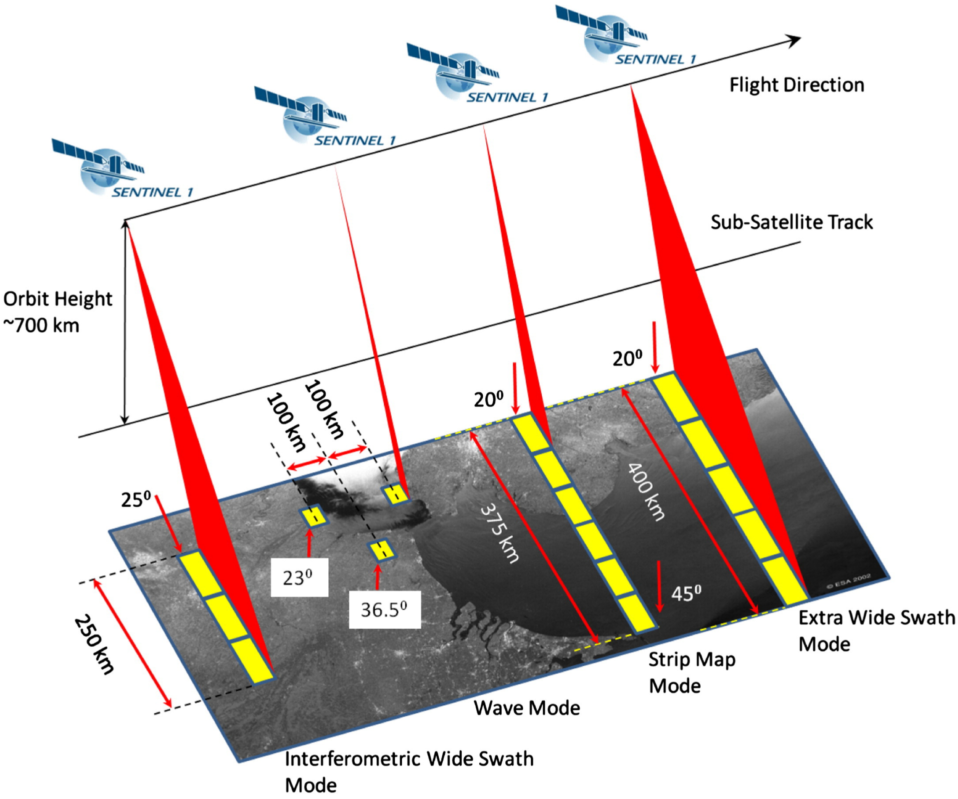
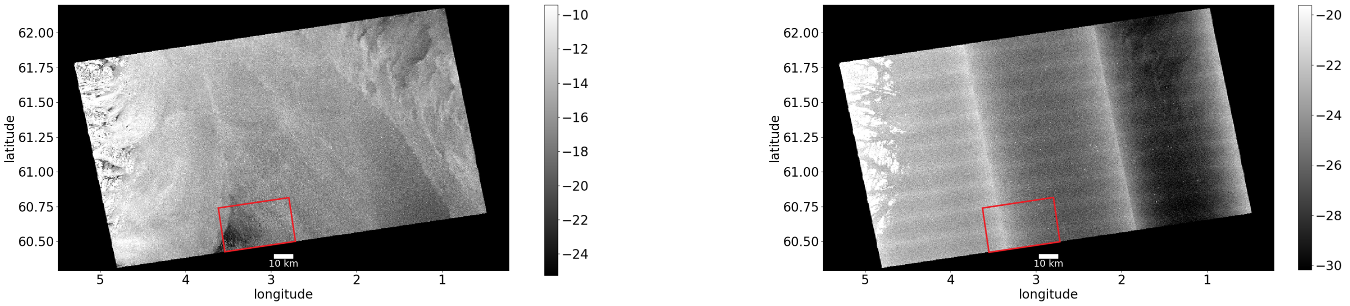


Figure 4.3: Overview of Sentinel-1 acquisition modes. Image used with permission from ESA. Image Copyright: ESA

The images used in this thesis are acquired in the IW swath mode, where the mode specific information is presented in Table 4.1. The polarization options for this data are dual-polarization HH+HV, VV+VH and single polarization HH, VV. HH+HV and single HH is only used over polar environments and sea-ice zones. In all other scenarios, Sentinel-1 employs VV+VH or VV. The data used in this thesis is hence in VV+VH mode. Figure 4.4 shows an example of the images used in this thesis.

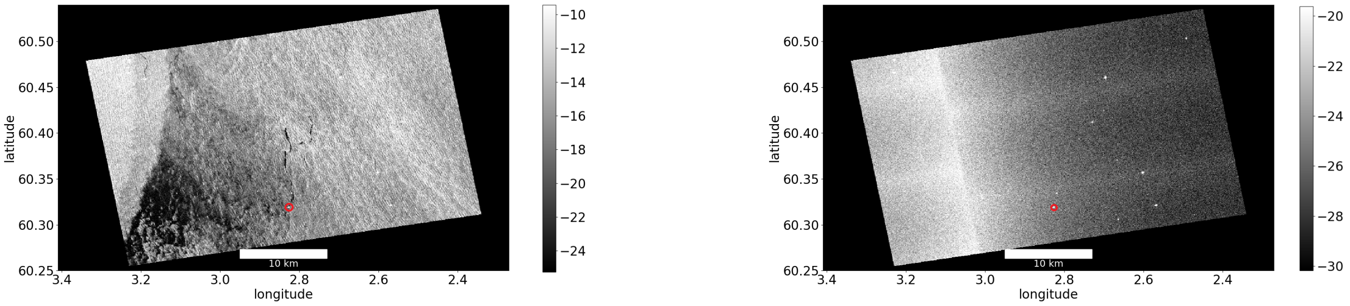
Table 4.1: Overview of some key characteristics for Sentinel-1 data in the IW mode.

Sentinel-1 IW properties	
Range resolution	20 m
Azimuth resolution	5 m
Altitude	700km
Global revisit time	6 days
Frequency	5.405 GHz
Polarizations	HH, VV, HH+HV, VV+VH
NESZ	-22 dB



(a) Whole scene in VV

(b) Whole scene in VH

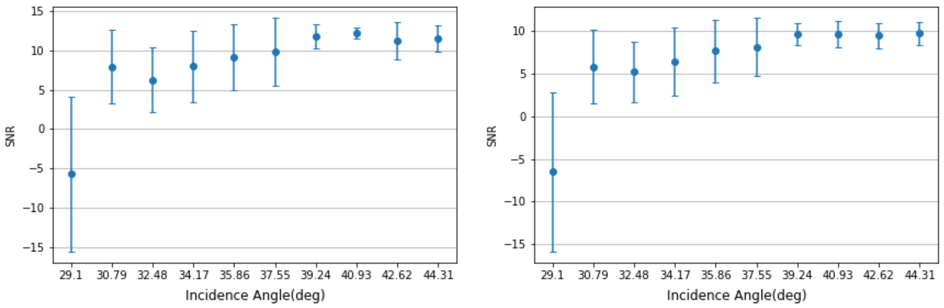


(c) Subset of scene around Brage (red circle) in VV

(d) Subset of scene around Brage (red circle) in VH

Figure 4.4: Example of a scene from the dataset. These images are displayed in decibel after speckle filtering and calibration to sigma nought. The scene covers the Brage platform and the subsets are a 40x40 km area, with the platform in the middle. Location of the subsets are located with a red rectangle in subfigure a and b. Brage is indicated by a red circle in subfigure c and d.

To evaluate the quality of the data we can examine the SNR. The average SNR for the scenes can be seen in figure 4.5 plotted against the incidence angle. As the figure shows, for most parts of the scenes the SNR lies around 10 dB. For the regions of interest, the slick areas, the SNR is lower and the average SNR for the produced water slicks is calculated to be between -5 and 5 dB in the VV band. This is unsuprising as the backscatter is lower for the slicks and with lower signal strength the SNR will be reduced.



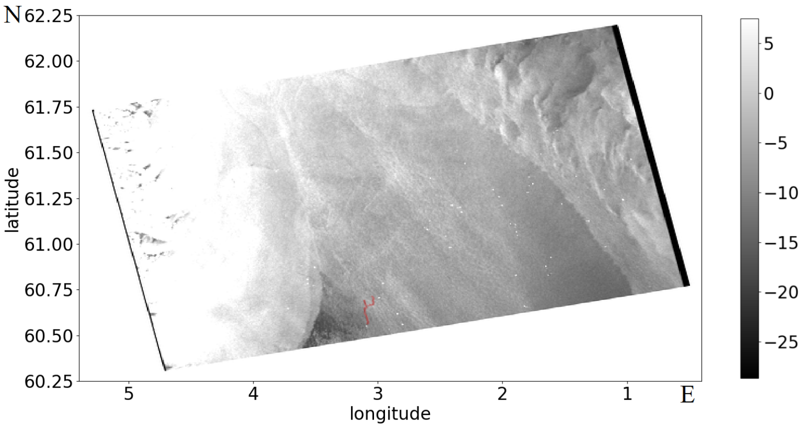
(a) VV **(b)** VH

Figure 4.5: SNR plotted against incidence angle. The mean SNR of all scenes in the dataset is shown by the dot and the bar illustrates the standard deviation.

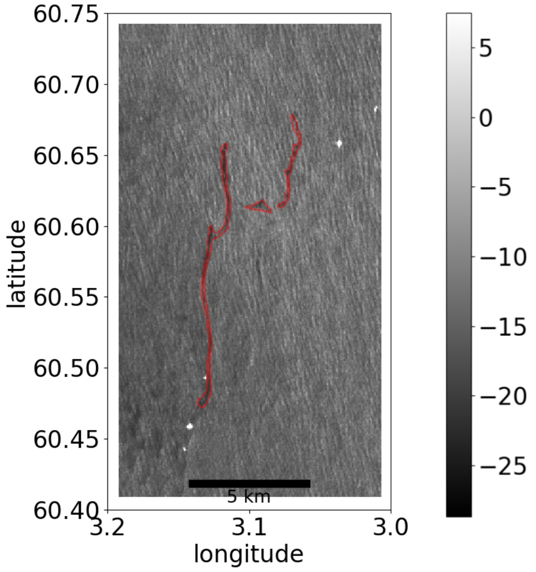
The 388 scenes in the dataset are separated into a training set and a validation or test set. The training set consists of 288 scenes, while the validation set consists of 100 scenes.

4.3 Ground Truth

Ground truth-based training and validation data is necessary to apply a machine learning method to the detection task. The oil spill detection services at Kongsberg Satellite Service (KSAT) (www.ksat.no) have graciously provided polygons with detections from the Sentinel-1 scenes in question to create training and validation data. This data consists of manually created outlines of all features in the images the experts at KSAT have deemed to be Produced water slicks. Figure 4.6 shows one of these polygons superimposed on the corresponding scene.



(a) Full scene with polygon



(b) Zoomed in on slick

Figure 4.6: Polygon marking a detected produced water slick plotted onto the corresponding scene. This image is of the Brage Platform located at about 60.5425∘N, 3.0468∘E.

KSAT makes detections based on a set of given parameters depending on the current task. For the North Sea and the Norwegian Sea detections they are based on size, length, and shape. For a slick visible in a scene to be categorised as a produced water slick it must be smaller than $5km^2$ and shorter than $5km$. It is also expected that a produced water slick is tail-like and connected to a platform. These parameters are the basis for the creation of the polygons made by the human operators. This information was gathered through personal communication with Ola Ørjavik, a KSAT operator, on the 11th of June 2021.

The polygons delineating slick observations were used to create binary masks to be used as ground truth for training data. The polygons were given as single polygon files, and detections from the same scene were separated. First, each polygon referring to the same scene was merged into one file, then they were then transformed into binary masks. These masks were geocoordinated with their scenes of origin and subsetted to cover the same areas as the scene. A subset of one of the scenes can be seen together with the corresponding mask in figure 4.7.

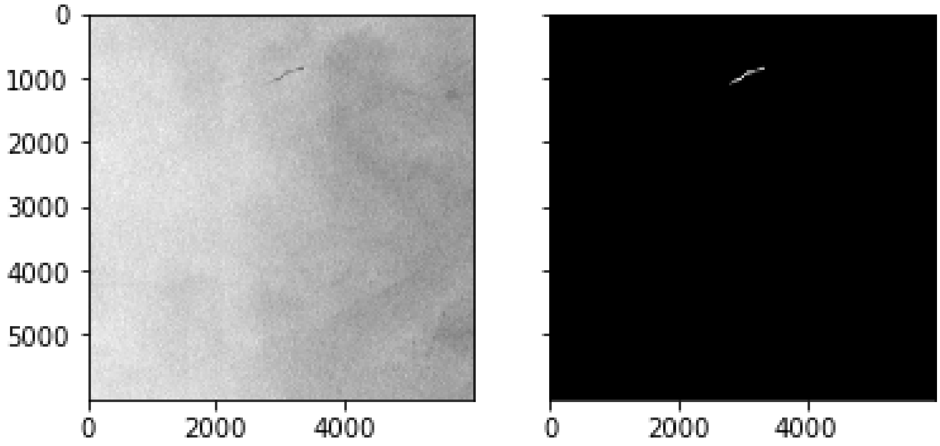


Figure 4.7: One scene over the Norne platform subsetted to 6000x6000 pixels on the left with the corresponding and subsetted mask of the produced water slick on the right.

These masks are for the work in this thesis considered ground truth. They are assumed to correspond with the actual oil slick detection and the delineations of these slicks. When training and testing the model the masks will be the reference data and it will also be the basis for evaluating the strengths and weaknesses of different approaches. The ground truth was also used before analysis to reduce the number of scenes processed. Any scene that is found to contain no slicks according to the reference data is ignored. It is worth noting that the platforms continuously release produced water into the surrounding waters when it is operational. This implies that the lack of detected slicks in some scenes is not in fact due to there being no oil in the waters around the platform, but the visibility of the slicks being too poor for the human operators to detect it. This is the reasoning behind excluding these images from the dataset, as they will provide no useful data to the deep learning neural networks. If the visibility of the slick is too low for human operators, the neural network based model can not be expected to detect it, and training on these images could contribute to confusing the model. The number of scenes in this dataset, 388, is a result of this selection process.

4.4 In Situ Data

It is of interest to consider the performance of the proposed approaches under different local conditions around the platforms. The two local conditions that is considered in this thesis are the wind conditions and the release data for the relevant platforms. In situ data from both platforms has been made available for this thesis and is used in chapter 7 and 8 to evaluate the effects different conditions might have on results. The in situ data consists of weather data and release information from the platforms.

The effects of oil concentration and produced water volume has also been discussed in [21]. The release data provides information on both volume of produced water released and the oil concentration in this produced water on a daily basis. The release data consists of daily averages for the oil concentration and the total produced water volume over the course of 24 hours. The concentration and volume can be multiplied to get the total oil released into the ocean. The releases do not vary much throughout the course of a day, so it can be assumed that daily release data is representative for the local conditions at the time the image is taken [21].

The operators of the two platforms, Wintershall for Brage and Equinor for Norne, have graciously provided release data for the periods examined in this thesis. Over the three year period our dataset was collected, Brage had an average concentration of oil in the disposed produced water of 15.56 g/m^3 and an average volume of discharged water of 14207 m^3, while Norne had an average concentration of 8.66 g/m^3 and an average volume of 19035 m^3

It has been shown that weather conditions can both effect the detectability of oil spills and be used to differentiate them from other phenomenon [59]. The weather data contains information on wind speeds that we know could effect the visibility of slicks in the SAR images. Wind data is divided into hourly observations. Figure 4.8 shows the daily wind speeds for the two platforms over the course of a year. We see here that in general, the average wind speed is a bit lower during the summer months. The wind data used in this thesis is acquired from the Norsk Klimaservicesenter (https://klimaservicesenter.no/).

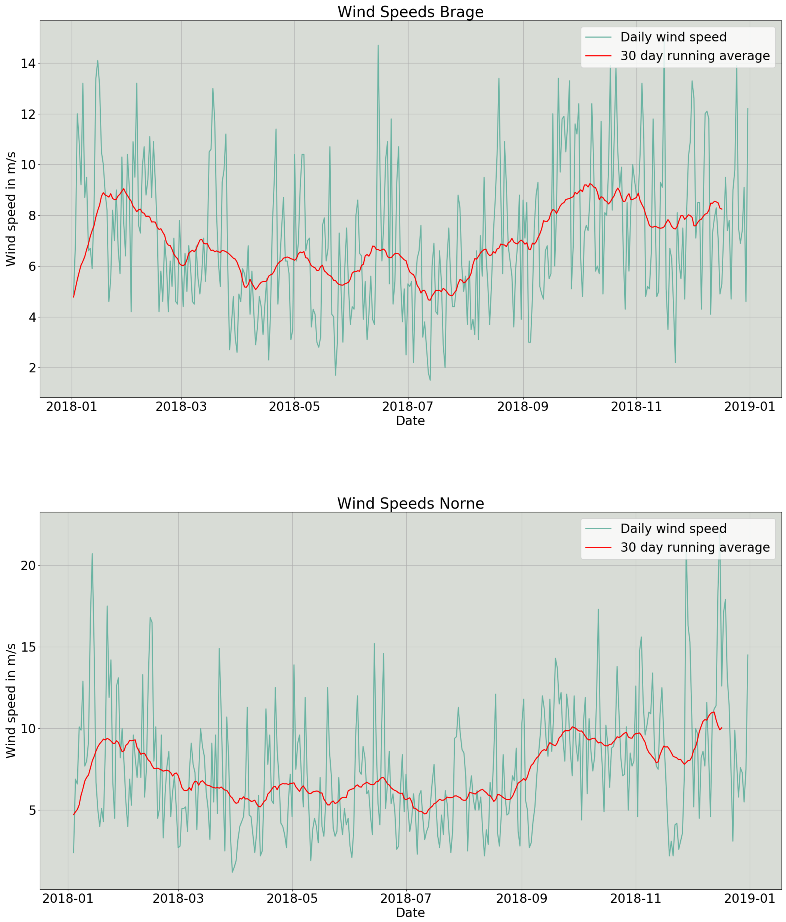


Figure 4.8: Daily wind speed at the two platforms over the course of a year.

5

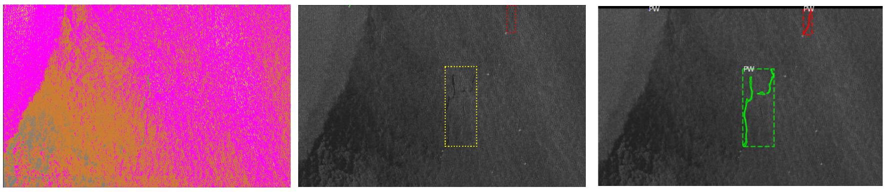
Theoretical Background for Method

In this chapter the concepts needed to understand the method applied are introduced. This includes the basics of image analysis and the segmentation algorithm and machine learning algorithm used in this thesis.

5.1 Image Analysis

The task of classifying different groups based on statistical attributes is one that has been worked on for hundreds of years [60], constanly being improved and applied to new fields. Since the advent of digital imagery, classification has been one of the most prominent types of image analysis. Pixel classification in digital images can broadly be divided into semantic segmentation, object detection and instance segmentation. Each of these have over the years been solved in many different ways [61–63]. We will not here go into detail for every approach to these tasks but will provide an overview of what each task entails.

Semantic segmentation refers to the process of dividing an image up based on the classes that are contained in it [64]. Each pixel is assigned to a segment based on some data or information and then grouped with the other pixels that are deemed to belong in the same segment. This process can include classifying each segment, but can also only consist of grouping the pixels in some meaningful way. Each segment is usually marked by a different color. Object detection is, as the name suggests, a process meant to find objects in an image of specific types. This consists of finding the objects location in the image and marking it somehow, usually with a bounding box. Instance segmentation combines the two other processes by first detecting objects in the image, then segmenting only the objects themselves. An example of the result for each of three process can be seen in figure 5.1. In the case of this work we will be applying instance segmentation since we are interested both in the location of slicks and their shape.



(a) Semantic Segmentation **(b)** Object Detection **(c)** Instance Segmentation

Figure 5.1: Three forms of image analysis done on the same scene from the dataset.

5.2 Segmentation Algorithm

For the work in this thesis, we will be applying the segmentation algorithm described here [37]. This is a method for semantic segmentation of marine SAR images. It is based on a Gaussian mixture model and takes into account the incidence angle degradation. Earlier methods have adjusted for the incidence angle effects in multiple ways, such as using Lambert's Cosine Law [36], the Minnaert Model [65], and variants of the Bragg Model [50]. All these approaches have in common that they assume a global degradation. This method, on the other hand, considers that different classes do not necessarily share incidence angle decays. The segmentation algorithm assumes a multivariate Gaussian distribution for the clusters along constant incidence angle azimuth lines.

$$p_{X,\Theta}(\mathbf{x},\theta) = \sum_{k=1}^{M} \pi_k \frac{1}{2\pi^{d/2}|\Sigma|^{1/2}} \exp(-\frac{1}{2}(\mathbf{x}-(\mathbf{a}_k-\mathbf{b}_k\theta))^T \Sigma_k^{-1}(\mathbf{x}-(\mathbf{a}_k-\mathbf{b}_k\theta))) \tag{5.1}$$

This is a Gaussian distribution where the expectation value is dependent on the incidence angle, which we have noted is assumed to be a linear relation in dB. $\mathbf{a}_k$ represents the mean value in decibels at $\theta = 0$ and $\mathbf{b}_k$ is the decay rate that is specific for the cluster k.

After the primary segmentation, the algorithm applies a Markov Random Field smoothing to improve homogeneity within the different segments. This is meant to avoid lone pixels that are segmented differently than the surrounding area. Markov Random Field (MRF) is a probabilistic model that, when used in image segmentation, is meant to capture an area's homogeneity by giving weight to the context of a given pixel. The conditional probability density function (cpdf) is calculated for a pixel and the cliques it is a part of consisting of its neighbors. We use the cpdf of a center pixel with intensity i_k given its neighbors $i_{k_1}, i_{k_2}, i_{k_3}, i_{k_4}$. Then from this we derive a cost function for the MRF $U(i_k, i_{k_{1...4}}) = V_{C_1}(i_k) + V_{C_2}(i_k, i_{k_{1...4}})$ where:

$$V_{C_1}(i_k) = 3ln[p_{I_k}(i_k)] \tag{5.2}$$

$$V_{C_2}(i_k, i_{k_{1...4}}) = \sum_{j=1}^{4} \frac{A(i_k, i_{k_j})}{B(i_k, i_{k_j})} - ln[J_0[\frac{C(i_k, i_{k_j})}{B(i_k, i_{k_j})}]] + ln[B(i_k, i_{k_j})] \tag{5.3}$$

Here $V_{C_1}(i_k)$ and $V_{C_2}(i_k, i_{k_{1...4}})$ are the single-clique and pair-clique potential functions. Each pixel is treated and placed into the segment that provides the highest cpdf. This can be done by minimizing the derived cost function.

This algorithm is designed to counteract the issues introduced by incidence angle degradation. During testing, it was observed that this effect was negligible after the images were reduced in size to only contain a 40x40 km area around the platform. This is in contrast to the whole image that is 250 km across in which the incidence angle degradation is significant and has a noticeable effect on segmentation. The incidence angle adjustment included in the algorithm might pick up on other differences in the across-track of the scenes, such as antenna noise. This could result in false adjustments for incidence angle degradation and corrupted results. In that case, it will be necessary to remove the component dealing with the incidence angle effect from the algorithm and instead use a simple gaussian mixture model.

$$p_{X,\Theta}(\mathbf{x}, \theta) = \sum_{k=1}^{M} \pi_k \frac{1}{2\pi^{d/2}|\Sigma|^{1/2}} \exp(-\frac{1}{2}(\mathbf{x} - (\mathbf{u}_k)^T \Sigma_k^{-1}(\mathbf{x} - (\mathbf{u}_k))$$

Integrated into the segmentation algorithm are options to do user-defined multilooking and reducing the number of pixels processed. Multilooking increases radiometric resolution while decreasing spatial resolution. Pixel reduction decreases the run time also at the cost of spatial resolution.

5.3 Mask Region-Based Convolutional Neural Network

The slicks that we are looking for in this work are often easily recognizable by the human eye because of their distinctive shape and the contextual information of their location in relation to oil platforms and other image features. However, it is often more difficult to statistically or mathematically distinguish the slicks from other areas and ocean surfaces in the images. This provides a challenge when it comes to automatically identifying the slicks without human arbitration.

Mask Region-Based Convolutional Neural Network (Mask R-CNN) is a deep neural network aimed at image segmentation and object detection [66]. Mask R-CNN is an expansion on the Faster R-CNN algorithm as described in [67] with the added benefit of creating a semantic mask for the detected object. It produces a bounding box containing the object instance and an object mask. It is capable of working on both videos and images. Mask R-CNN is currently viewed as the state-of-the-art approach to neural net-based object detection and instance segmentation [68]. It has been applied to many different tasks and shown viability for a wide variety of applications [68–70].

Mask R-CNN works in two stages. The first stage proposes an area of the image where an object of interest might be located using a bounding box. The second stage is the classification of the detected objects and creating a mask to describe the objects' location in the image more precisely. These stages are connected to the backbone structure, Feature Pyramid Networks (FPN) [71]. FPN consists of a bottom-up pathway, a top-bottom pathway, and lateral connections. The Bottom-up pathway is a convolutional neural net that extracts features from the images. The top-bottom pathway creates feature pyramid maps that are similar in size to the bottom-up pathway. The lateral connections are adding operations and convolutional operations between the two pathways on corresponding levels. The advantage of FPN is that it maintains strong semantic features at differing resolution scales.

First, a Mask R-CNN produces proposals for regions that may contain objects using a region proposal network (RPN) [72] that scans all FPN top-bottom pathways. This also requires the use of anchors that bind features to their location in the original image. The anchors in a Mask R-CNN are boxes with predefined locations and scales relative to the image. The anchors bind to different levels of the FPN top-bottom pathways and uses their predefined position to propose likely locations of objects in the image. This information is then used to check the areas of interest more closely, and guarantees that the proposed locations are not moved as the image is processed.

The second stage consists of another neural net that generates object classes, bounding boxes, and masks. These are aligned to the original image using information provided by the RPN in the first stage of the process. We will only be working with two classes for this work; produced water and background, and will only be evaluating the accuracy of the delineation based on the resulting produced water masks. This can be expanded to include more classes, and the bounding boxes can be used as a faster way to get detections without the delineations, but this is outside the scope of this thesis. Figure 5.2 visualizes how Mask R-CNN functions.

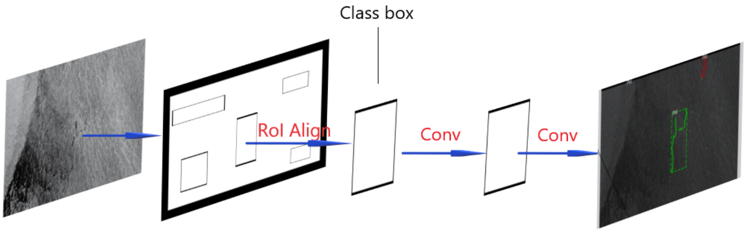


Figure 5.2: Mask RCNN first proposes possible regions of interest (RoI) based on anchors. Of these regions of interest, any above a certainty threshold is selected and the object within that bounding box is classified. Then the region of interest is run through two convolutional networks to create a mask of the instance.

As a supervised machine learning process, Mask R-CNN is dependent on training data. In this thesis, the training data will be binary masks of produced water detections based on the observations provided by KSAT. The algorithm used for this thesis was built around Matterports implementation of Mask R-CNN [73].

6

Methodology

In this chapter an overview of the methods applied to the task of detection and delineation of produced water slicks in this thesis is presented. The work in this thesis consists of two stages. The first is detection of the produced water slicks in the scenes. This includes both finding low backscatter areas and the identification of these areas that represent slicks and not look-alikes. The second is the delineation the produced water slicks. The level of success for the whole process will be measured by a combination of the methods' ability to detect the produced water slicks and the classification accuracy. Three different approaches are proposed here based on the two algorithms described in section 5.2 and section 5.3. Figure 6.1 shows a flowchart of the workflow of the method in this thesis.

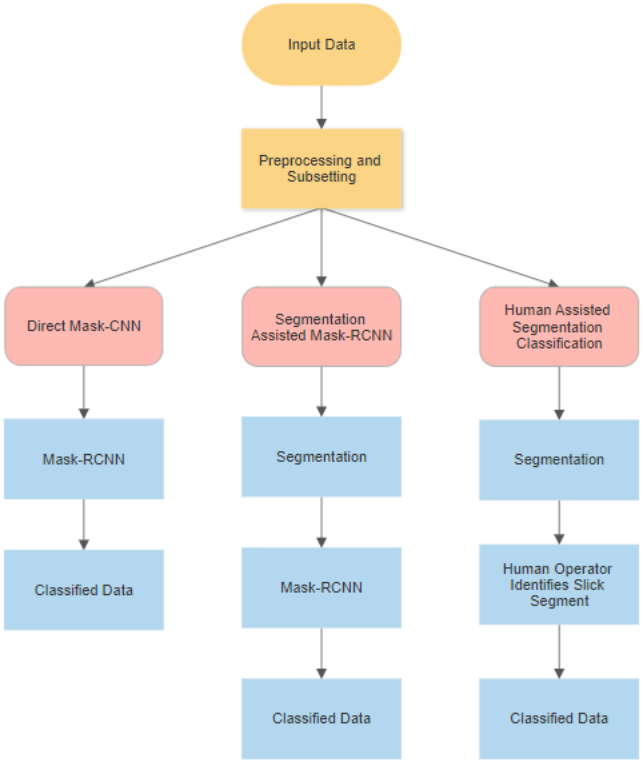


Figure 6.1: Flowchart of the different approaches used in this thesis. The yellow boxes are the common steps for all approaches. The red boxes represent the names of each individual approach and the blue boxes are the specific steps for each approach.

6.1 Preprocessing Data

Because of difficulties in differentiating the slicks in the VH-band, it has been excluded from the analysed dataset. Within this work the results are derived using the VV-band. The SAR scenes are first subsetted to approximately 40×40 km around the two relevant platforms to reduce data size and computation times. Thereafter is the data calibrated using the provided meta data [30], then a 7×7 Lee Sigma speckle filter is applied to remove speckle noise. Finally, the data is converted to decibel for better analysis and visual representation. An overview of the preprocessing steps is shown in figure **??**.

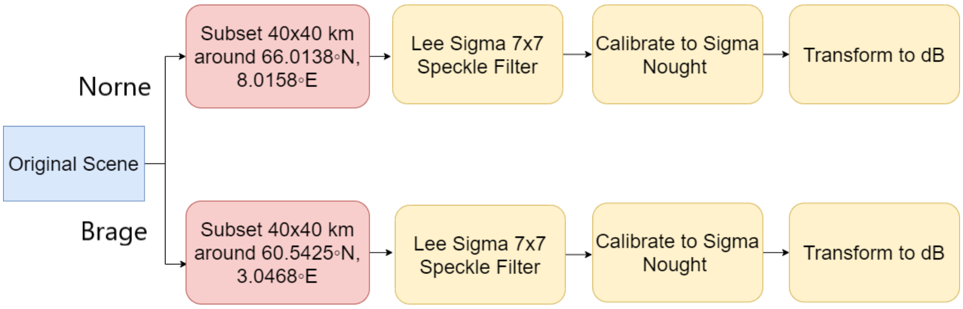


Figure 6.2: Flowchart of preprocessing steps.

6.2 Approaches

Here we propose three alternative methods for detection and delineation of produced water slicks in SAR images. These methods will be compared against each other for detection and delineation accuracy. Each method automates the process either entirely or partly and is meant to be more efficient alternatives to the existing process of manual detection by human operators.

6.2.1 Direct Mask Region-Based Convolutional Neural Network

In the Direct Mask Region-Based Convolutional Neural Network (DMRCNN) approach the Mask R-CNN algorithm is applied directly to the images after the preprocessing stage. The images are processed as described in section 5.3. First, the method is trained on the training set consisting of 288 scenes and subsequently validated using the remaining 100 scenes to evaluate performance and calculate the precision of the model. The training data consists of SAR images, processed in the same way as described in the pre-processing stage, and their corresponding masks made from the KSAT polygons. An example of the training data used can be seen in figure 6.3.

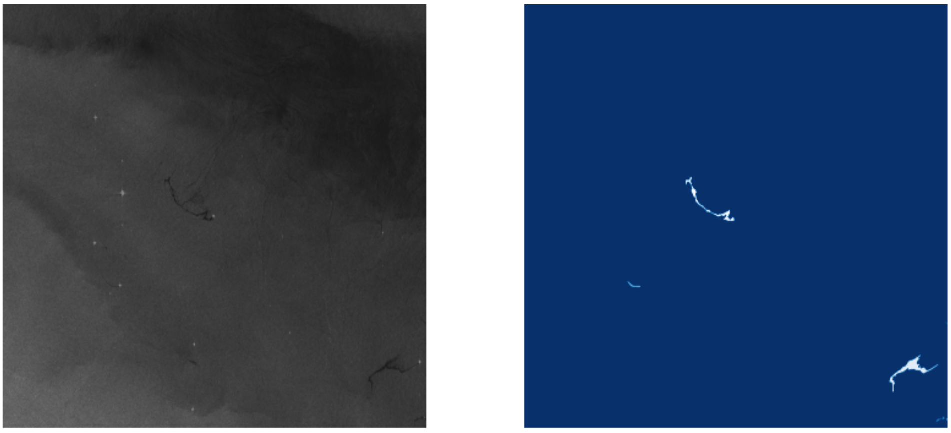


Figure 6.3: An example of the training data used for DMRCNN. The data consists of the subsetted and preprocessed scenes and masks for the instances in the scene. The image to the left is the image as fed to the neural network, the image to the right is the masks used as reference data.

The Mask R-CNN was fed training data consisting of the preprocessed and subsetted scenes, a bounding box containing the detections and the binary mask.Table 6.1 shows the training parameters for the model used in this thesis. The training time for this model was 42 hours.

Table 6.1: List containing the specific parameters used for training Mask R-CNN in this thesis. The parameters were chosen based on experimentation and trial and error.

Parameters for Training of Neural Network

Epochs	10 epochs for training head layers + 40 epochs for training all layers
Learning Rate	0.001 for first 10 epochs, 0.0001 for last 40 epochs
RPN Anchor Scales	(16, 32, 64, 128, 256)
Steps Per Epoch	300
Validations steps	50
Regions of Interest per Scene	10
Backbone	Resnet101
Minimum Detection Confidence	0.8

The algorithm will use all information contained in the dataset to make predictions about detection and delineation. This includes size, spatial positioning, context, and backscatter intensity. The information gathered during training is used to create weights that are then used to predict detections in new images.

The training data is passed through a series of layers as any convolutional neural network to create the model.

Directly applying the Mask R-CNN method to the dataset should give the algorithm access to all the information in the images. This would give the method a broad set of features to create a model for object detection. The Mask R-CNN has been shown to be promising for general oil detection in earlier research [74] when applied directly to a dataset, and this approach also had the advantage of simplicity. It only needs to be fed the relevant data without any other analysis performed on it and is fully automatic and not reliant on human intervention.

6.2.2 Human Assisted Segmentation Classification

Human Assisted Segmentation Classification (HASC) is a semi-automatic approach and is divided into two stages. First, an image is segmented using the segmentation algorithm described in section 5.2. We can see an example of the resulting product in figure 6.4. The slick is visible in the segmented image, but the slick is segmented together with the low wind areas in the image. It is therefore necessary to employ another approach to differentiate it. For that reason, the resulting product is passed over to a human operator that identifies the produced water slick in the image and marks it. The marked and segmented area containing the slick is then used to generate a mask or polygon of the observation.

This approach combines the segmentation algorithm with the traditional detection done by a human operator. This is intended to limit the time used by the operator to search for possible slicks in scenes by first passing it through the semantic segmentation of the algorithm. This will possibly provide more precise delineation than what a human operator can, given the processing of individual pixels that the segmentation provides. One of the most significant weaknesses of the segmentation algorithm is that it only does its segmentation based on backscatter values, not on any spatial parameters or context clues. By adding the benefit of a human operator, one could filter out false detections and misclassified areas and even potentially differentiate look-alikes and different kinds of oil slicks. Though not fully automatic and still demanding human intervention, this method would automate the delineation and simplify detection. This could prove a viable alternative to a wholly automatic approach.

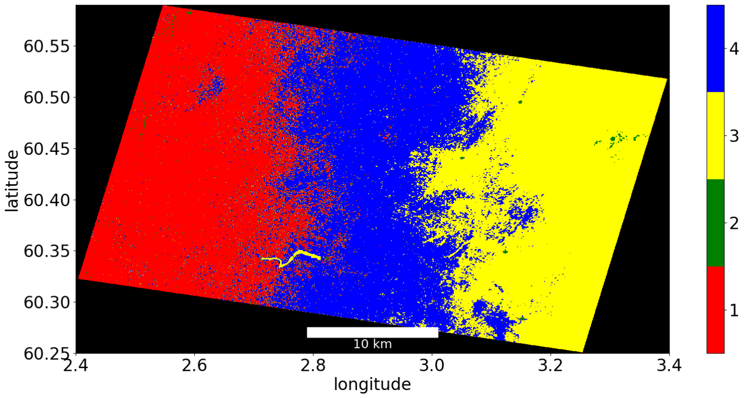


Figure 6.4: A scene containing a slick from the Brage platform segmented by the segmentation algorithm. The slick has been segmented together with the low wind area in the right part of the image, but is clearly distinguishable for a human operator. Each colors represent one segment as given by the colorbar. These segments have not been classified.

Unlike the other approaches, HASC does not need training data only a validation set for testing. Because this approach is not automatic the validation set for this method consists of only ten random scenes from the whole validation set.

6.2.3 Segmentation Assisted Mask Region-Based Convolutional Neural Network

Segmentation Assisted Mask Region-Based Convolutional Neural Network (SAMR-CNN) combines probabilistic semantic segmentation with Mask R-CNN. Mask R-CNN is applied as described previously, but the input data is the segmented images produced by the algorithm outlined in section 5.2 and not the original scenes. An example of the training data used for this approach can be seen in figure 6.5. The process will consist of two parts. First, the segmentation algorithm is applied to the images to create segments such as the one shown in figure 6.4. Then those segments are passed to the Mask R-CNN algorithm as training data together with the binary masks. The Mask R-CNN processes the training data and creates a model for object detection and delineation of the slicks. Then the model is used to identify and delineate slicks in a different set of segmented images to evaluate performance.

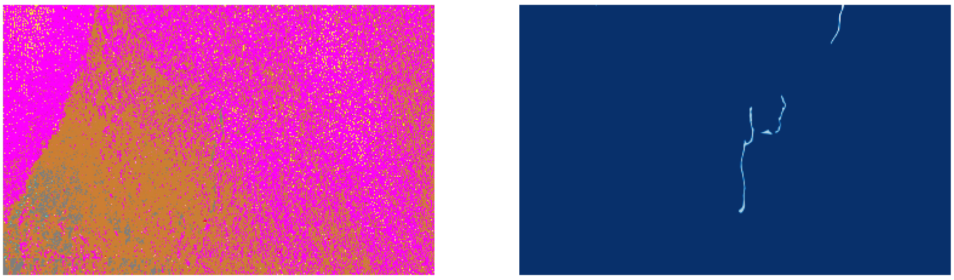


Figure 6.5: Training data used for Segmentation Assisted Mask R-CNN. The data consists of the subsetted and segmented scene and masks for the instances in the scene. The image to the left is the image as fed to the neural network, the image to the left is the masks used as reference data.

By first performing a semantic segmentation of the images based on pixel intensity, one could more clearly divide up potential areas containing slicks from surrounding water surfaces. When processing this, the neural net will have fewer parameters to base its model on, focusing on spatial, pattern, and shape information. Though more demanding than the direct Mask R-CNN approach, it could prove to perform better. This approach is more time demanding in the detection phase which could prove problematic for implementation for a practical application.

For this approach the Mask R-CNN was fed training data created from segmented scenes. The model was trained on a training set consisting of 288 scenes and their corresponding masks and tested on a validation set consisting of 100 scenes and their corresponding masks. Table 6.2 shows the training parameters for the model used in this thesis. The training time for this model was 37 hours.

Table 6.2: List containing the specific parameters used for training Mask R-CNN in this thesis. The parameters were chosen based on experimentation and trial and error.

Parameters for Training of Neural Network

Epochs	10 epochs for training head layers + 35 epochs for training all layers
Learning Rate	0.001 for first 10 epochs, 0.0001 for last 35 epochs
RPN Anchor Scales	(16, 32, 64, 128, 256)
Steps Per Epoch	300
Validations steps	50
Regions of Interest per Scene	10
Backbone	Resnet101
Minimum Detection Confidence	0.8

6.3 Evaluation

For it to be possible to draw any conclusions on the quality of the approaches proposed in the thesis, a set of criteria for evaluation must be established. It must be concluded exactly what is looked for in the results produced from each approach. All objective measures of the results will be based on the reference data that consists of the polygons from KSAT. These will be regarded as ground truth when evaluating the performance of the approaches.

The first objective for each approach is to detect the slicks in the scenes. To evaluate this performance the detection rate of slicks will be used. This will be calculated by the number of correctly identified slicks, independent of the precision of the slicks delineation, divided by the number of slicks in the scenes according to the reference data. In addition the number of false positives will be recorded. False positives will be the number of slicks identified by the approach that are not slicks according to the reference data divided by the total number of slicks identified. To evaluate the delineation the masks of the slicks created by the approaches will be compared with the masks of the reference data at the pixel-level. Omission errors and commission errors are both of interest when considering the quality of the delineations, thus both the average producer's accuracy and the average user's accuracy will be presented for each method and used to evaluate performance [1]. Producer's accuracy is the number of pixels correctly classified in a class by the classifier divided by the number of

---

[1] Producer's accuracy is also known as recall and user's accuracy is also known as precision.

pixels that are in that class according to the reference data. The user's accuracy is the number of pixels classified as the given class by the classifier divided by the total number of pixels in that class according to the classifier.

$$a_u = \frac{N_c}{N_{pred}} \tag{6.1}$$

$$a_p = \frac{N_c}{N_{ref}} \tag{6.2}$$

Where a_u is the user's accuracy, a_p is the producer's accuracy, N_c is the number of pixels correctly classified as a given class by the classifier and N_{pred} and N_{ref} are the number of pixels classified in that class by the classification method and the reference data consisting of KSAT polygons respectively. Another measure of accuracy often used in machine learning is the *F1 score*. This measure takes into account both the user's accuracy and the producer's accuracy to provide a complete measure of the performance of a model. The F1 score for each approach will also be presented in chapter 7.

For the purposes of practical application, the training time for the machine learning methods is less relevant than the time used for predictions after a model is trained. The time cost for each approach will therefore be the time it takes each approach to produce a result from a new scene presented to it. The in situ data from the platforms will in this thesis be used to evaluate the flexibility and robustness of the approaches under differing local conditions related to wind and release data. As the goal is to provide a more efficient way to detect and delineate produced water slicks the time cost and workload for each method will also be discussed. A visual inspection and subjective evaluation will also be presented to make note of trends and aspects not obvious in the numeric data.

$$F1 = 2\frac{a_u a_p}{a_u + a_p} \tag{6.3}$$

Criteria for Evaluation:

- Detection rate
- False positives
- Producer’s accuracy
- User’s accuracy
- F1 score
- Time cost
- Flexibility with differing local conditions
- Subjective visual inspection

In addition to looking into the overall performance of the approaches, each approach is also tested on ten random scenes from the dataset from here on referred to as the experimentation set. This is to better visualize how each of them perform on specific scenes under specific conditions. These scenes were randomly selected using a digital random number generator and sorted by date.

7

Results

In this section we present the results of experimentation with the approaches described in section 6.2. DMRCNN, SAMRCNN and HASC have each been tested on Sentinel-1 marine SAR images taken over the Brage and Norne platforms. The results are meant to give an indication of each methods ability to detect and delineate produced water slicks. The evaluation of the experimentation will be based on the criteria as stated in section 6.3.

The five measurable criteria for performance have been recorded and presented in table 7.1. Here the accuracies are the average for each approach for the slicks that were detected by it. This means that any slick that the approach failed to detect does not count toward this accuracy. The detection rate here is the percentage of all slicks in the validation set the approach was able to detect. We can see from table 7.1 that DMRCNN has a significantly higher detection rate than the two other approaches. SAMRCNN and HASC do on the other hand achieve more precise delineations than DMRCNN. DMRCNN has the lowest time cost and is significantly faster than both SAMRCNN and HASC. DMRCNN produces the most false positives and HASC the least with zero for all scenes it was tested on.

Table 7.1: Results for the all approaches. Detection rate signifies the percentage of slicks detected by the approach. The producer's and user's accuracy are the average for each over all scenes. Both accuracies are based on the slicks that were detected by the approach. The time cost is the time it takes the approach to process one scene and produce a result. These results are based on the whole validation set consisting of 100 scenes for DMRCNN and SAMRCNN and the experimentation set (10 scenes) for HASC.

Overall Results			
	DMRCNN	**SAMRCNN**	**HASC**
Detection Rate	81%	53%	50%
False Positives	15%	11%	0%
Producer's Accuracy	81%	87%	88%
User's Accuracy	76%	81%	91%
F1 score	64%	44%	45%
Time Cost	5.7 Seconds	13.4 Minutes	11.2 Minutes

While the overall results provide useful information on its own, its is also of interest to discuss how the different approaches compare on individual scenes. Table 7.2 contains an overview of the accuracy for the experimentation set as well as the average accuracy for all scenes in validation set. The accuracy here is given by the producer's accuracy and includes the instances where an approach was unable to detect a slick.

Table 7.2: Summary of results for all three approaches on the experimentation set. For each scene the table includes the date the SAR image was taken, the average wind speed (WS), the average oil concentration (Conc) and the total released produced water volume (Vol) that day. The release data is only for the named platform, and therefore only represents one slick in the image. The accuracy given are the producer's accuracy for each scene. The average is the average accuracy for each method over all scenes, including the ones where it was not able to detect slicks.

Date	**Platform**	**WS (m/s)**	**Conc (g/m^3)**	**Vol (m^3)**	**DMRCNN**	**SAMRCNN**	**HASC**
02.04.18	Brage	4.3	15.00	10135	76%	88%	92%
15.08.18	Brage	8.2	18.46	19285	64%	71%	74%
20.12.18	Norne	7.4	5.1	25462	74%	–	–
08.02.19	Norne	4.1	11.79	20437	79%	86%	89%
06.06.19	Norne	5.9	10.74	17594	80%	87%	88%
26.11.19	Brage	2.0	19.69	16203	–	–	–
25.02.20	Norne	5.7	5.5	21484	73%	–	–
17.06.20	Brage	4.2	21.42	8864	82%	88%	90%
17.08.20	Brage	2.6	8.4	12176	79%	–	–
01.10.20	Norne	5.5	16	20776	83%	84%	83%
Average Whole Validation Set					**66%**	**55%**	**57%**

7.1 Direct Mask Region-Based Convolutional Neural Network

This approach was able to detect 81% of all slicks in the scenes, and 15% of detections produced by the model were false positives. The average producer's accuracy was 81% and the average user's accuracy 76%. The trained model takes on average 5.7 seconds to process one scene. Table 7.3 shows the results of DMRCNN for the experimentation set. The detections produced by this model do, in general terms, match the ground truth as established by the human operators of KSAT. There is some inaccuracy with precisely identifying the edges of the slicks and a tendency of confusion with long, narrow slicks and slicks containing complex geometries. In some cases the model will also identify two parts of the same slick as separate entities, particularly in long and narrow slicks. This is for example the case for the scene dated 02.04.2018 in table 7.3. Here one long slick is divided into to separate detections, causing the model to detect three slicks in a scene that in reality only contains two. In figure 7.1 we can see an example of the results of the DMRCNN. The delineations in this seem to be good based on a visual inspection, but we see that for the purple slick it cuts out when the slick grows too narrow and for the turquoise only a very small part of the slick is delineated.

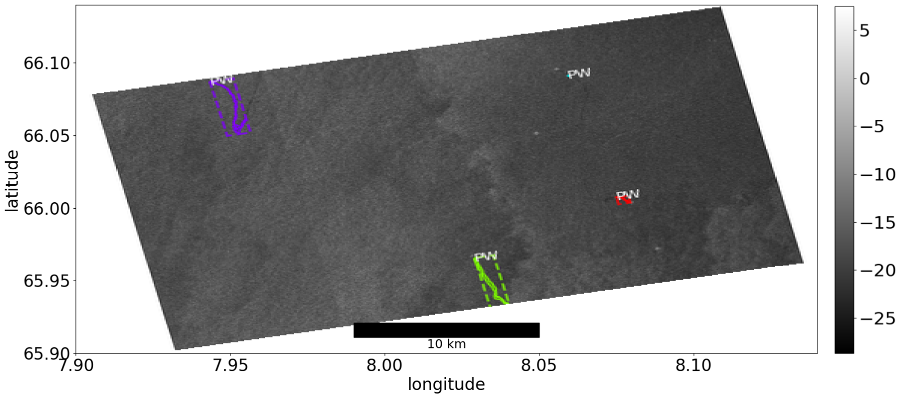


Figure 7.1: One result produced by the Direct Mask R-CNN approach. The rectangles made up of dotted lines are the bounding boxes of each detection and the colored shapes within them represent the mask of the slick as predicted by the method.

Table 7.3: Results for Direct Mask R-CNN on the experimentation set. For each scene the table includes the date the SAR image was taken, the average wind speed, the average oil concentration and the total released produced water that day. The release data is only for the named platform, and therefore only represents one slick in the image. Also included are the number of slicks in the scene, the number of slicks detected by the method and the producer's accuracy for each scene.

Date	Platform	WS (m/s)	Conc (g/m^3)	Vol (m^3)	Slicks	Detected	Producer's (Accuracy)
02.04.18	Brage	4.3	15.00	10135	2	3	76%
15.08.18	Brage	8.2	18.46	19285	4	3	64%
20.12.18	Norne	7.4	5.1	25462	1	1	74%
08.02.19	Norne	4.1	11.79	20437	1	1	79%
06.06.19	Norne	5.9	10.74	17594	1	1	80%
26.11.19	Brage	2.0	19.69	16203	7	0	–
25.02.20	Norne	5.7	5.5	21484	1	2	73%
17.06.20	Brage	4.2	21.42	8864	1	1	82%
17.08.20	Brage	2.6	8.4	12176	1	1	79%
01.10.20	Norne	5.5	16	20776	1	1	83%

7.2 Segmentation Assisted Mask Region-Based Convolutional Neural Network

SAMRCNN was able to detect 53% of all slicks in the scenes and 11% of detected slicks were false positives. For the scenes where the segmentation was able to segment out the slicks, the end result after running the segmented scene through the model was an average producer's accuracy of 87% and an average user's accuracy of 81%. In cases where the method was not able to detect slicks it was most often caused by the slick being segmented together with the surrounding waters by the segmentation algorithm. This approach used an average of 13.4 minutes to produce detections. In general the detection for this approach was underwhelming, but the delineation once a slick was detected was closely in line with the reference data as illustrated by the producer's and user's accuracies. An example of a result from this method can be seen in figure 7.2. The delineation seems very precise, but a false positive is also included in the lower part of the image.

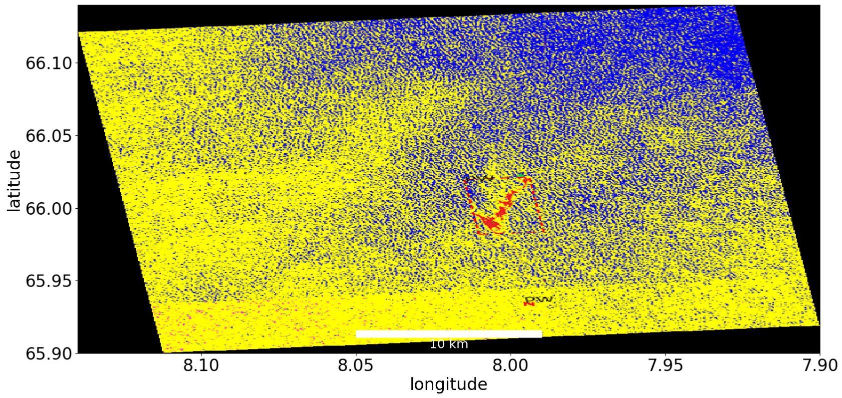


Figure 7.2: One result produced by the Segmentation Assisted Mask R-CNN approach. The rectangles made up of dotted lines are the bounding boxes of each detection and the colored shapes within them represent the mask of the slick as predicted by the method

Table 7.4: Results for Segmentation Assisted Mask R-CNN on the experimentation set. For each scene the table includes the date the SAR image was taken, the average wind speed, the average oil concentration and the total released produced water that day. The release data is only for the named platform, and therefore only represents one slick in the image. Also included are the number of slicks in the scene, the number of slicks detected by the method and the producer's accuracy for each scene.

Date	**Platform**	**WS (m/s)**	**Conc (g/m^3)**	**Vol (m^3)**	**Slicks**	**Detected**	**Accuracy**
02.04.18	Brage	4.3	15.00	10135	2	2	88%
15.08.18	Brage	8.2	18.46	19285	4	4	85%
20.12.18	Norne	7.4	5.1	25462	1	0	–
08.02.19	Norne	4.1	11.79	20437	1	1	86%
06.06.19	Norne	5.9	10.74	17594	1	1	87%
26.11.19	Brage	2.0	19.69	16203	7	0	–
25.02.20	Norne	5.7	5.5	21484	1	0	–
17.06.20	Brage	4.2	21.42	8864	1	1	88%
17.08.20	Brage	2.6	8.4	12176	1	0	–
01.10.20	Norne	5.5	16	20776	1	1	84%

7.3 Human Assisted Segmentation Classification

This method was tested on the experimentation set. In six out of the ten, slicks were clearly separated from the surrounding water and could be picked out. In the other four the slicks were segmented together with the surrounding water and therefore impossible for a human operator to identify. There were all together twenty slicks in the scenes of which ten were detected. This results in a detection rate of 50%. In the scenes where the slick was possible to pick out the approach produced an average users accuracy of 91% and an average producers accuracy of 88%. Omission errors were more common than commission errors meaning that compared to the ground truth data, it is more common to make an error of excluding a pixel that is part of the slick than including one that is not. This could be because the polygons that are used to form the ground truth cover a slightly larger area than the slick itself because of slightly inaccurate delineation from the human operator. Table 7.5 shows the results for HASC for the experimentation set. In figure 7.3 an example of one of the resulting segments is shown.

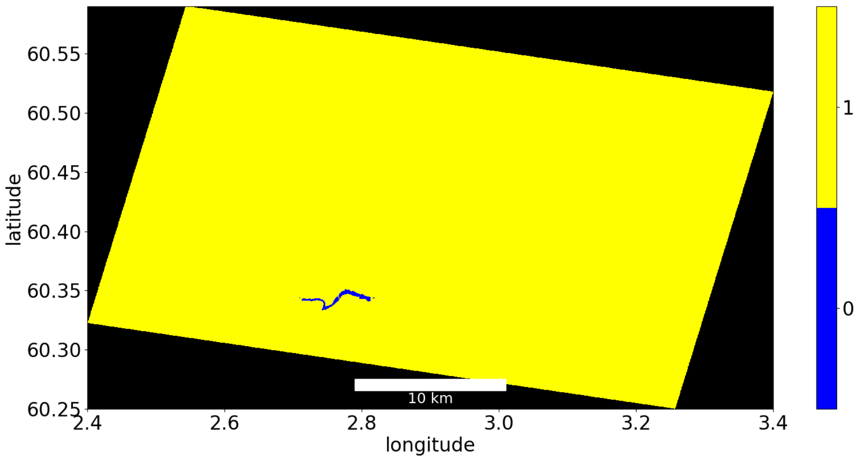


Figure 7.3: Result of one scene classified done by human assisted segmentation classification. The yellow pixels are classified as background and the blue pixels are classified as produced water.

Table 7.5: Results for Human Assisted Segmentation Classification on the experimentation set. For each scene the table includes the date the SAR image was taken, the average wind speed, the average oil concentration and the total released produced water that day. The release data is only for the named platform, and therefore only represents one slick in the image. Also included are the number of slicks in the scene, the number of slicks detected by the method and the producer's accuracy for each scene.

Date	Platform	WS (m/s)	Conc (g/m^3)	Vol (m^3)	Slicks	Detected	Accuracy
02.04.18	Brage	4.3	15.00	10135	2	2	92%
15.08.18	Brage	8.2	18.46	19285	4	4	89%
20.12.18	Norne	7.4	5.1	25462	1	0	–
08.02.19	Norne	4.1	11.79	20437	1	1	89%
06.06.19	Norne	5.9	10.74	17594	1	1	88%
26.11.19	Brage	2.0	19.69	16203	7	0	–
25.02.20	Norne	5.7	5.5	21484	1	0	–
17.06.20	Brage	4.2	21.42	8864	1	1	90%
17.08.20	Brage	2.6	8.4	12176	1	0	–
01.10.20	Norne	5.5	16	20776	1	1	83%

7.4 Effects of Weather Conditions and Oil Release

To evaluate the performance of each approach under different conditions the dataset was organized based on in situ data. To get a large enough dataset both the training set and the validation set was used for these experiments. For the first experiment the dataset was grouped by local wind speeds around the given platform, and the results of this can be seen in table 7.6. In general the data shows that very high an very low wind speeds make the slicks difficult to detect, which is to be expected as both have earlier been shown to lessen the damping ratio of the slicks. This effect clearly has a stronger impact on the segmentation-based approaches.

Table 7.6: Producer's accuracy of the approaches under different wind conditions. The wind speeds are the average wind speed for each day. The accuracies are the average accuracies of how many slick pixels were correctly classified. For HASC ten randomly selected scenes from each group was used.

Wind Speed(m/s)	DMRCNN	SAMRCNN	HASC	Average
0.0-2.0	27%	0%	0%	9%
2.0-4.0	67%	21%	23%	37%
4.0-6.0	77%	59%	63%	66%
6.0-8.0	81%	63%	69%	71%
8.0-10.0	79%	66%	67%	71%
10.0-12.0	64%	43%	45%	51%
>12.0	17%	4%	6%	12%

For the second experiment the scenes were organized by the daily oil release from the given platform. The oil release was calculated by multiplying the oil concentration of the produced water with the released produced water volume resulting in the total oil released over the course of a day. The results of this experiment can be seen in table 7.7. In general it was shown that increased oil release led to better detection. This is to be expected as increased oil release leads to larger and more visible slicks as we can see from figure 7.4. It is clear from this plot that there is some correlation between the slick size and the released oil from the platform.

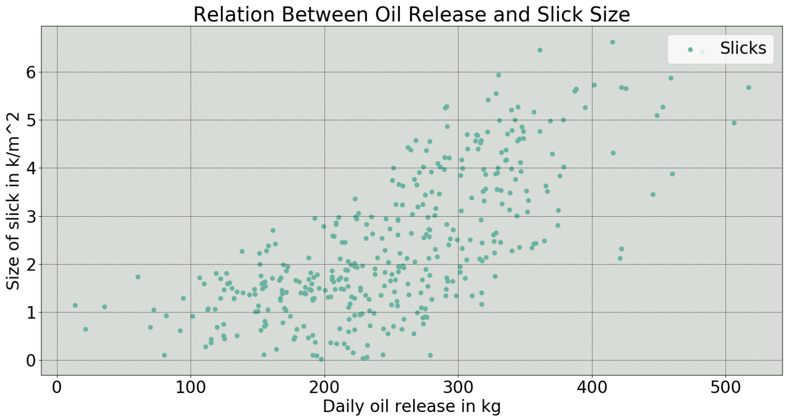


Figure 7.4: Scatterplot of all slicks plotted with oil release for the given day and the size of the slick as determined by the reference data. The daily oil release is calculated by multiplying the average oil concentration of the produced water with the daily release volume.

Mask R-CNN was shown to have some issues when delineating very large slicks which explains the dip in accuracy for the two approaches that employs deep learning when the oil content gets very high. Through a visual inspection in the cases where Mask R-CNN underperformed it seems that in particular long and narrow slicks or slicks with complex geometries produced lower detection rates and higher degrees of inaccuracy for both approaches that employ Mask R-CNN. One such slick is illustrated in figure 7.5.

Table 7.7: Producer's accuracy of the approaches in relation to the amount of oil released from the platform. The Oil Content is the daily release from the platforms calculated by multiplying the concetration of the produced water with the produced water volume. The accuracies are the average accuracies of how many slick pixels were correctly classified. For HASC ten randomly selected scenes from each group was used.

Oil Content(kg)	DMRCNN	SAMRCNN	HASC	Average
<100	32%	35%	46%	38%
100-150	59%	47%	51%	51%
150-200	71%	54%	56%	60%
200-250	77%	59%	61%	66%
250-300	83%	62%	63%	69%
350-400	79%	67%	69%	72%
>400	74%	65%	72%	70%

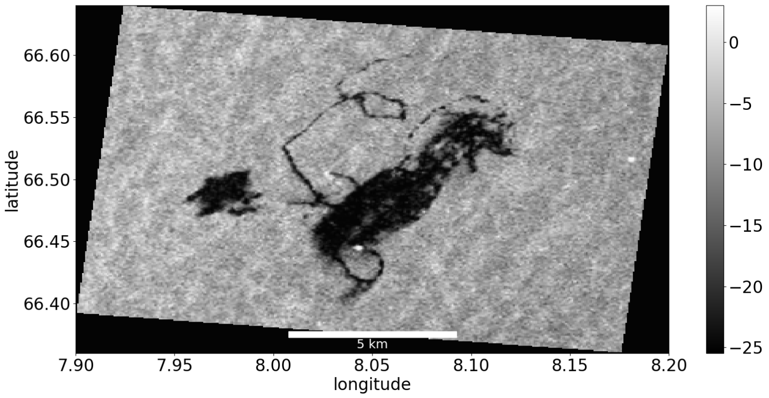


Figure 7.5: A scene of complex geometric shapes around the Norne platform.

8

Discussion

In the previous chapter the experimental results for each of the proposed approaches were presented. Here we discuss the results and their implications. First, it is of interest to compare the approaches in accordance to the stated goal of creating and evaluating robust approaches to produced water slick detection and delineation. Thereafter the effects of external factors are discussed before some error sources for this work are presented.

8.1 Comparison

HASC was shown to have very clear delineations by it's high producer's and user's accuracy in the cases where it detected a slick. Through visual inspection and comparison with the original scenes this approach did even in some cases seem more precise than the reference data produced by human operators. HASC did however have a severe drawback in it's ability to detect slicks, with around five out of every ten slick being impossible to identify in the segmented image. The segmentation algorithm is slow, taking upwards of ten to fifteen minutes to produce a segmented image. This high time cost is especially problematic when working with large amounts of data. HASC has the disadvantage, when compared to the other approaches, that it is not fully automatic. It is still dependent on human input to produce classifications. Even so it does produce very accurate delineations with a very low degree of human involvement. The forced human involvement could also function as a fail-safe and guarantee the quality of the results.

SAMRCNN is the slowest of the approaches tested in this thesis. It consists of two steps, wherein the first segmentation takes upward of ten to fifteen minutes as pointed out earlier. This becomes a bottleneck, resulting in long processing times, even though the neural network is reasonably fast. SAMRCNN is fully automatic and produces very precise delineations, but is hampered by the low rate of detection in cases where the segmentation is unable to distinguish slicks from the surrounding waters. There are also examples of the neural network not being able to detect a slick even if the segmentation algorithm successfully separates it, but this issue is far less common. As with HASC, the segmentation algorithm too often segments slicks in with the background, making it impossible for the trained model to identify them. When the slick is segmented successfully by the segmentation algorithm, this approach represents a fully automatic method for detection and delineation of produced water slicks with more precise delineations than DMRCNN as made clear by the accuracy of SAMRCNN in the cases where it is able to detect a slick. We can see from table 7.4 that the producer's accuracy is very high in the cases where a correct detection is made, but the overall performance of the approach is reduced by the low detection rate.

DMRCNN, when trained, is fast compared to the other approaches, producing results in only a few seconds. It can process and produce both bounding boxes and masks for produced water slicks in the scenes efficiently. Using the trained model is also simple and can be run on large datasets simultaneously. The network can be trained for different datasets if the scenes one need to observe somehow differs greatly from the ones used in this thesis. The Mask R-CNN models does however take a long time to train. Depending on the training parameters, training times of between 30 and 45 hours were recorded. Even though the Mask R-CNN can be trained for differing data, this is time demanding. DMRCNN was shown to have the highest detection rate of all three approaches tested in this thesis. The delineations produced were not as precise as the two other approaches and suffered particularly for slicks above 15 kilometers in length and less than 100 meters across and slicks containing complex geometric shapes.

Both approaches that employ Mask R-CNN exhibited reduced accuracy and detection for long narrow slicks and slicks with complex or unique geometries. These types of slicks are not very common and are mostly formed when the oil released through produced water over the course of a day is above around 400 kg. The rarity of these types of slicks is the most likely explanation for the reduced accuracy of the machine learning based approaches. These are dependent on training on the object they are meant to identify, and it is expected that rare features will be more difficult for them. In earlier works *complexity* (C) and *spreading* (S) have been used to describe the shape aspects of the slicks [20]. Average accuracy seems to start decreasing at a threshold value for both S and C. When either S becomes too low or C becomes too high there is a drop-off both in delineation accuracy and rates of detection. The dataset used here is not large enough to pinpoint these exact threshold values or the relation between these attributes and the accuracy, but it is still worth noting this weakness in these two approaches. Increasing the size of the dataset could both provide the ability to map this relation and contribute to reduce it's effect. A larger dataset can be assumed to contain more examples of both long narrow slicks and slicks with complex geometries, giving the Mask R-CNN model more experience with these types of slicks. Increasing the number of epochs for the training stage of the model was also noted to decrease this discrepancy but did result in overfitting of the model and a decreased accuracy for the validation dataset.

$$C = \frac{P}{2\sqrt{\pi A}} \tag{8.1}$$

$$S = \frac{100\lambda_2}{\lambda_1 + \lambda_2} \tag{8.2}$$

Where P is the perimeter or length in kilometres of the slick, A is the area of the slick and λ_1 and λ_2 are the two eigenvalues associated with the computed covariance matrix.

Because of the higher detection rate of DMRCNN it in sum produces the best result of the approaches described in this thesis. HASC and SAMRCNN do however produce more precise delineations with HASC having the most precise delineations when taking into account only the scenes where the approach was able to detect the slicks. HASC also produces no false positives. Of the three only DMRCNN and SAMRCNN are fully automatic, whereas HASC depends on a human operator. SAMRCNN was the slowest to produce detections with HASC being only slightly faster. DMRCNN was by far the fastest of the approaches when trained. Given the large number of platforms on the norwegian continental shelf, the amount of data produced by satellites and the need for fast production of results, the time needed to make detections, may be one of the most significant factors when evaluating practical viability. DMRCNN performs better overall as shown by it's F1 score of 64% compared to 44% and 45% for SAMRCNN and HASC respectively. SAMRCNN and HASC do however, produce more accurate delineations in the scenes where slicks are detected. Going through individual results for this approach shows that the largest hindrance for detection is the segmentation algorithm. This implies that an improved method for semantic segmentation could drastically improve results for both HASC and SAMRCNN.

A comparison could be made to the results achieved in similar project to this thesis [15]. When attempting to automatically detect oil spill in SAR images using deep learning, it was reported a F1 score of 73.1% for the best hyperparameters. The best approach proposed in this thesis, DMRCNN, had an f1 score of 64%. The discrepancy between these two perfomances is most likely a result of the differing subject matters. While produced water is a type of oil slick, it is often characterized by being smaller than the oil spills used in [15] and also usually creates very thin films [11]. These factors make produced water spills more difficult to detect than spills as discussed in section 3.2. As [15] concludes with this method being practically viable based on these results one could argue that with the improvements that will be discussed in section 10.1 one or more of the proposed approaches in this thesis could also be viable.

8.2 Effects of Weather Conditions and Produced Water Releases

Environmental conditions have some clear implications on the performance of all the proposed methods in this thesis. The produced water slicks are visible due to the dampening of capillary and gravitational waves caused by the surface film as discussed in section 3.2. The wind has an effect on the damping ratio of a slick as theoretically discussed in [75]. This was also shown empirically in this thesis and has also been shown in earlier research [50, 55]. For very low wind speeds the damping ratio becomes small since the surrounding waters become smoother. Comparing the SAR images with the in-situ data, a minimum threshold for the slicks being observable seems to be around wind speeds of 2 m/s. When the wind speed is too high the damping ratio also becomes low, because the dampening effect of the film is not sufficient to counteract the waves caused by the wind. The upper limit for detectable slicks is around 11.5 m/s. These observations also closely resemble the conclusions reached in other studies [21, 59].

The two approaches that employ a deep neural network, DMRCNN and SAMRCNN, are in general more successful at detecting larger slicks than smaller ones. In both cases it is far more common for slicks that cover few pixels to be ignored by the model, although the accuracy of the delineation is in general better for smaller slicks once the slick has been detected. The data does imply that the amount of oil released affects slick size as seen in figure 7.4. With higher oil releases, either through increased release of produced water or higher oil concentration in the produced water, the slicks in general become larger. Conversely this implies that oil concentration and volume of produced water has an effect on the detectability of the slicks for the deep neural network based methods. Higher volumes of oil released results in improved detection for these two approaches. This does however not hold as true for HASC where the ability to distinguish the slicks is mostly dependent on the backscatter and not the size of the slick, given that the algorithm processes the scenes at the pixel-level. The segmenation algorithm provides functionality to process fewer pixels by using multi-looking or mere pixel reduction. This makes the algorithm produce results faster but also result in not every pixel being processed individually. In the case of reduced pixel count for increased segmentation speed, some slicks are too thin and may not appear in the segmented image. Some very small or narrow slicks also seem to disappear as a result of the smoothing that is done to maintain homogeneity of segments in the image. This is the most likely explanation for there being any effect at all from oil releases on HASC.

Local conditions that have not been studied in this work that might have an effect on results include rain and snow and ocean currents. It has been suggested in earlier work [76] that heavy rainfall might contribute to quicker dispersion of oil slicks due to the turbulence this produces on the ocean surface. The effect was especially prominent for light oil releases which the produced water slicks discussed in this thesis would be classified as. This implies that heavy rainfall would shorten the time a slick would be observable within the images, meaning the likelihood of detection could go down since constant observation is not available. Ocean currents are known to be one of the main drivers of an oil slicks movements together with wind drag [77]. This means that both the direction of the slick and it's spread is partially dictated by the local ocean currents. With the knowledge that particularly the DMRCNN approach's performance is dependent on the size and form of the slick, one can conclude that ocean currents could potentially have a large effect on a slick's detectability.

8.3 Error Sources

At the border of an object or between two different classes in an image one can often find mixed pixels. These are pixels that do not obviously belong to either class and often have a pixel value that lies somewhere in between the typical values for the two classes [78]. This is most often a result of the surface area covered by the pixel containing both types of surfaces making the resulting intensity a combination of the two. This becomes an issue when doing image segmentation and semantic object detection because the class affiliation of the pixel is ambiguous. When the data contains these ambiguities delineation becomes less precise and can contribute to lowering the performance of the object detection.

All the approaches proposed in this thesis classify every pixel binarily. This results in the issues related to mixed pixels and set an upper limit to the precision of delineations. Any produced water mask suggested by these methods can only be as precise as the pixels, meaning that the resolution of the image, in this case 5x20 meters, represents the maximum precision of the models ability do depict the ground truth.

Any supervised machine learning method is limited by the quality of the training data. In the case presented in this thesis, one limiting factor for training data accuracy are the potential human errors in the making of the detection polygons. These are based only on human observation of the SAR images. The delineation in some cases is sub-optimal as can for example be seen in figure 8.1. In this figure one can clearly see that the polygon does not perfectly align with the outline of the slick. This imperfection in turn makes the model created by the neural network less accurate in its delineation and therefore affects it's performance. A perfect approach could in this case at the very best be producing results that match exactly the ones created by human operators.

The imperfect nature of the polygons also have an effect on the evaluation of the result, since our empiric evaluation of the classification use the polygons as reference data. A method performing better than the capabilities of a human operator would by this evaluation standard have a lower accuracy than if the method matched the quality of human interpretation perfectly. DMRCNN, in particular, has a tendency to produce false positives and this is one of the weaknesses of this approach. When viewing these false positives, some seem to be actual slicks that the human operator has failed to detect, but without reference data to document this, the exact prominence of this occurrence is not possible to determine. These false positives are considered errors in the evaluation of the method, but could in reality represent an advantage compared to the current human-based slick detection work.

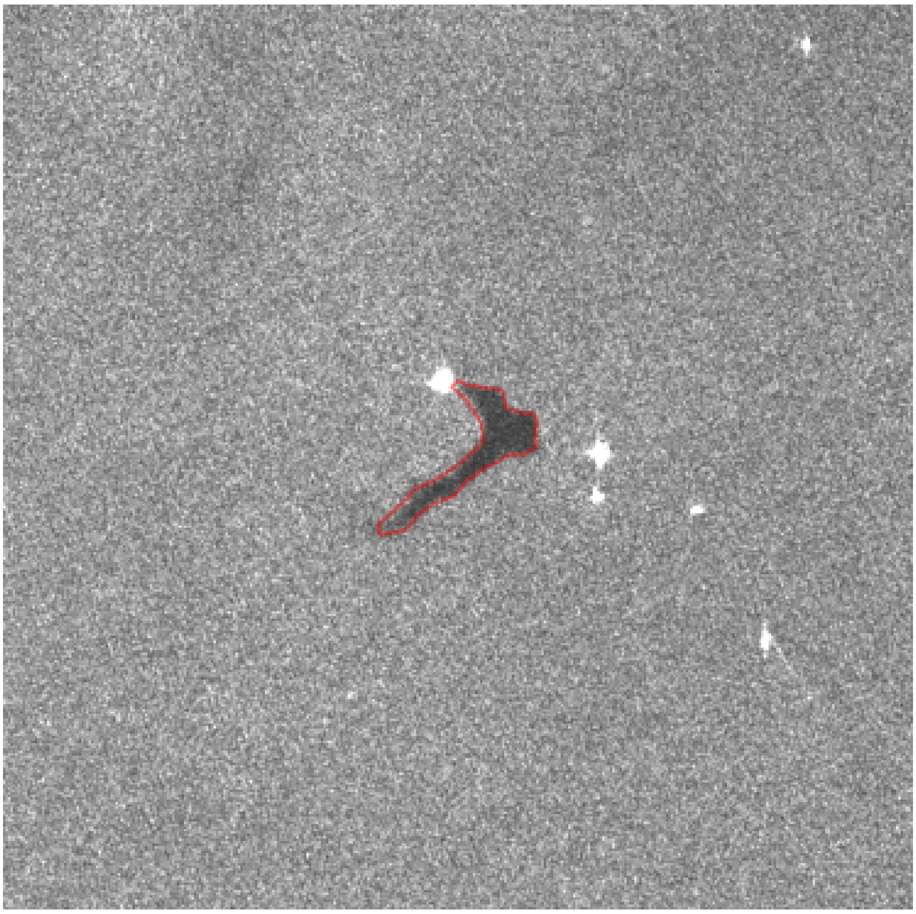


Figure 8.1: A polygon marking a produced water slick from the Norne platform.

The methods described in this thesis were designed based on a very homogeneous dataset. The scenes themselves were taken over only two areas and the variety in surface types included is limited. Both the training data and the validation data for the deep learning based methods were from this same dataset. It has therefore neither been tested nor trained on scenes from a different source or from a different area. This might result in the proposed method performing worse when presented with data from different areas. The Mask R-CNN algorithm is designed to counteract this somewhat by augmenting the data during training, but before the method is tested on a broader variety of different datasets the robustness and flexibility of the method can not be confirmed. In addition to this the dataset does not contain many examples of other types of slicks and look-alikes. This means that the methods ability to distinguish produced water slicks from other similar features may be limited.

The wind speeds used in this thesis to evaluate the effects local wind conditions have on detectability are based on hourly measurements at approximately the time the images were taken. For the purposes of a general understanding of the effects this is sufficient, but it is worth noting that the data used does not indicate the exact conditions at the exact time the SAR image was taken. As wind speeds can change drastically over very short time increments, the resulting analysis might be partially corrupted in cases where the current wind conditions varies greatly from the closest measurement. In addition to this, the measurements of the wind speeds used were taken at the platforms themselves. The local wind speed could vary throughout the image, and this can in fact be seen in some cases. This could effect the detectability, especially for slicks from other platforms in the scenes and this could in turn have an effect on the calculated results. These two factors combined lend some uncertainty to the evaluation of the effects of wind on a slicks detectability.

For HASC the human interpretation used for testing in this thesis has not been carried out by a trained expert. This might have affected the results for this approach in particular as a trained operator may produce better results than what is presented here.

9

Conclusion

The goal of this thesis was to create robust and flexible approaches for produced water slick detection and delineation. The models should be capable of detecting and precisely delineating the slicks while ignoring other features such as natural films and low-wind areas. The criteria for the work presented were stated as follows:

- Propose automatic or semiautomatic approaches for detecting and delineating produced water slicks in marine SAR images.
- Evaluate environmental and operational conditions that affect the performance of the approaches and the viability of automatic and semiautomatic methods of produced water detection.
- Identify future improvements to data or modelling techniques that could improve automatic detection and delineation of produced water slicks.

This thesis proposed three approaches to produced water slick detection and delineation in marine SAR images. Two of these methods are wholly automatic and based on deep learning with a convolutional neural network. The third is a semiautomatic approach meant to simplify and expedite the work for human operators. The neural networks were trained on pairs consisting of a Sentinel-1 SAR image and a binary mask drawn at the oil spill detection service at KSAT. The binary masks were regarded as ground truth and used for training and validation. Through extensive experimental work the methods were evaluated and compared based on both the ability to detect slicks and ignore other features in the scenes and the accuracy of the delineation. The three approaches proposed were Direct Mask R-CNN (DMRCNN), Segmentation Assisted Mask R-CNN(SAMRCNN) and Human Assisted Segmentation Classification(HASC).

Of the three approaches, DMRCNN was the most successful method when it came to detecting slicks in the scenes. This method is also the most easily implementable consisting of only a trained neural network that can be applied directly to SAR scenes. The network then performs both detection and delineation simultaneously. This approach did however have a slightly less accurate delineation than the two other approaches.

SAMRCNN combined the segmentation of the algorithm described in section 5.2 with Mask R-CNN. The workload was divided between the two in such a way that the segmentation algorithm performed the delineation and then the neural network performed detection on the segmented images. This did improve upon the delineation compared to the DMRCNN. This approach does however suffer in the cases where the segmentation algorithm is not able to distinguish the slicks in the scene from the surrounding waters. This approach is also more computationally and time demanding than the direct approach.

HASC resulted in the highest accuracy scores of the three approaches proposed in this thesis for the slicks it was able to detect. It did however have a detection rate of only 50%. Although not fully automatic like the other two approaches, it does represent a simplification of the process of detecting and delineating the slicks. The approach contributed highly accurate delineation in the cases where an instance was detected, but had the same issue as the Segmentation Assisted Mask R-CNN when it came to detection. More slicks were not detected than was the case for DMRCNN.

The limitations and effects of local weather conditions and release circumstances have been outlined in this thesis. It was observed that the local weather conditions and the concentration and volume of the released produced water had a significant effect on the results achieved. In general wind speeds below 2 m/s and above 11.5 m/s resulted in slicks being difficult to separate from the open water or difficult to detect. In particular the two approaches that employ the segmentation algorithm were sensitive to the wind conditions. The oil concentration and volume of released produced water has a significant effect on the size of a given slick. Mask R-CNN in particular performed worse when attempting to detect very small slicks with areas of around $1km^2$, resulting in there being correlation between the amount of oil released and the results of DMRCNN and SAMRCNN.

In sum the DMRCNN showed the most viability for use in a practical application. This approach had the highest accuracy when it came to detection of slicks, and is also the simplest to implement after training the model. It should however be noted that with improved semantic segmentation, SAMRCNN might have the largest potential for improvement.

10

Suggestions for Future Work

To make the approaches described in this thesis more viable for practical applications, improvements could be made. Here we have identified possible areas for further work that could strengthen the methods and improve the effectiveness of the classification and outlined some possible applications of this theoretical and experimental work.

10.1 Improvements

The dataset used in this thesis consists of dual-polarization data. Each scene includes as VV-band and a VH-band. As discussed in chapter 2 using multiple polarizations in a dataset can provide supplementary information and in that way improve the quality of image analysis. Adding dimensions to the data does however provide additional noise and particularly if the added band has a low SNR, such as is the case for the VH-band. Such bands provide very little useful information while providing a large amount of additional noise. The difficulty in interpreting the data in the VH-band is clearly seen in 10.1

Although the VH-channel was ignored in this work, research has shown that cross-polarization is viable for detection of man-made phenomenon in marine

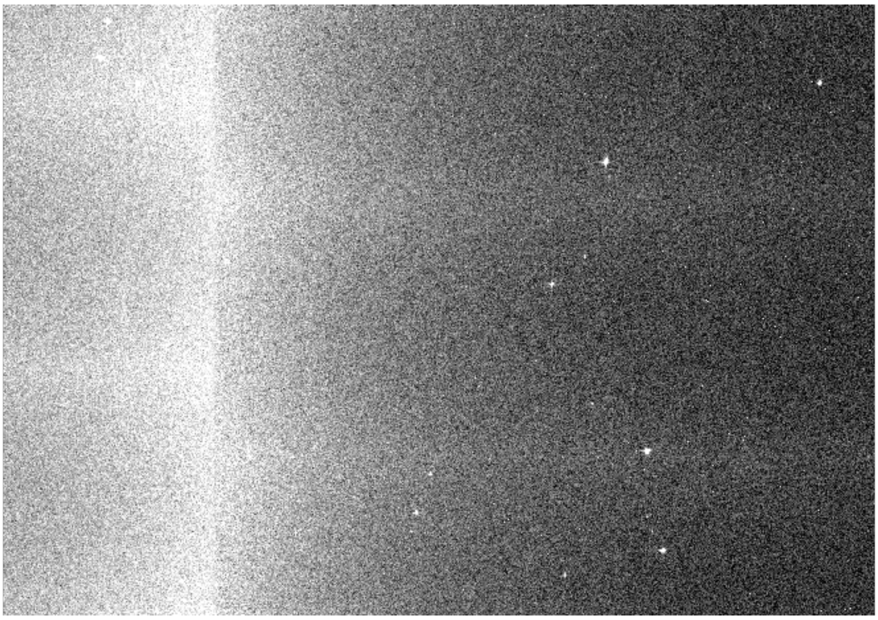


Figure 10.1: Subset of scene containing the Brage platform after calibration and speckle filtering presented in decibel.

areas [79]. These have very high backscatter in cross-polarization channels and we can also observe in figure 10.1 how bright ships and platforms appear. Since the produced water slicks that we are trying to detect in this research always originates from a platform, combining platform detection with the produced water slick detection presented in this thesis might improve performance. Platform locations could be used as an internal sanity check for the location of the slick in the algorithm or be included in the weights of the model. For future work it would be of interest to see how incorporating this information could effect the performance of the models described in this thesis.

In this thesis we have limited our dataset to consist of the Sentinel-1 scenes, the detections from KSAT and some in-situ data. This limitation was specifically chosen to work directly on detecting produced water slicks in the images without the use of other known variables. However, for practical applications there is no reason for excluding any information that can increases the efficiency and performance of the detection method. The exact geographic locations of the platforms is known and a usable metric for the method [20]. Instead of either detecting the platforms or including their location only as a bright spot in the image, their exact location could be fed directly into the algorithm either before detection or as a sanity check after the fact. By establishing beforehand the location of either all platforms in a given dataset or only the ones that are of interest one could exclude any slick or look-alike that does not originate from the platforms. This could also be used to ignore produced water slicks from other platforms if one is only interested in one or a few specific ones.

The Mask R-CNN that is utilized in this thesis is often referred to the state-of-the-art algorithm for object detection and instance segmentation [68]. It is built on top of the previous Faster R-CNN [67] algorithm and has been used for a wide variety of applications such as self-driving cars, medical imagery and microscopy [80–82]. It has proven to have good results for a myriad of these applications and different types of images. In this thesis we discuss the viability for the application on detection and delineation of produced water slicks. Since Mask R-CNN's creation in 2017 there has been some progress on improving on it [83, 84]. By using an improved version of the network in the future one could expect to achieve better results than what was noted in this thesis. Especially would improving the networks ability to detect very small objects significantly improve on the results. One proposed way to improve detection of smaller objects is to multiply the number of small objects in the training set [85]. This addresses the issue of small objects taking up relatively small parts of the images giving the network less data to train on and also results in more instances of small objects in the training set. As Mask M-RCNN detects individual instances this should not corrupt the data in any way and could prove a viable improvement of the network for the purposes of produced water detection.

When evaluating the effects of local weather conditions on the performance of the approaches in this thesis, only the current wind conditions were considered. Inertia in the local system, including surface waves and conditions and wind-driven currents would imply that the effects of wind are not completely instantaneous. It is likely that earlier wind conditions also have an effect on the directions of the slick and its detectability. In figure 3.2 for example, we can see that the slick is travelling in a different direction than the wind direction. Including existing work on these long term effects of local conditions and oil slick travel could provide interesting insight into the movement and drift of produced water slicks and potentially improve detection accuracy for the approaches proposed here [86].

10.2 Applications

As of now the main method used for detecting and delineating produced water slicks in SAR images is manual observation by human operators. Automating this process would cut manpower hours needed for these operations, potentially saving operating entities expenses and allow for more efficient use of human assets. To be viable for this use, an implementation must be built based on the theoretic background described in this thesis. Building on this work to create a program that automatically takes in SAR data as it's coming in, detects and delineates the images and stores it in a practical matter for the operating entity is a suggested area for further work. For the direct Mask R-CNN approach this would entail connecting the trained model to the existing framework of a given operator. Then running the images through the preprocessing steps and the model to produce results. For the Segmentation Assisted Mask R-CNN one has to create an implementation that first runs the image through the segmentation algorithm and then takes the result as input into the pretrained Mask R-CNN. Human Assisted Segmentation classification requires an implementation of the segmentation algorithm and a user interface for the human input. This user input should be fast and only need simple keyboard inputs. When this approach has been used in experiments for this thesis, the human input has been given through coding, which is not practical for real world implementations.

One of the main concerns for the monitoring of produced water releases is discrimintating between legal releases as part of regular operations and illegal releases. As of yet the main method for bringing to light illegal releases or determining the release volumes of oil and gas installations is on site measurements. This is a time demanding procedure that can be easily circumvented by groups operating in bad faith. Discerning information about the hydrocarbon content and volume of releases using remote sensing techniques would greatly improve governing entities ability to monitor the relevant operators. It could prove very useful to combine a statistical method [16] that can identify properties of the released produced water from SAR data with the approaches for automatic detection and delineation proposed in this thesis. This would provide daily automatic tracking of release data for every installation under a governing agency's jurisdiction.

An interesting application for the approaches proposed in this thesis arises from its potential synergy with inroads that have been made regarding detection of other ocean features [13, 15, 87, 88]. By combining automatic detection of produced water with automatic detection of other ocean features one could in theory create a method for automatic classification of all marine features from SAR data. A constant and wholly automatic classification of all features in marine SAR images would be useful for a wide specter of different fields such as military, environment and climate studies. This could also have significant academic and economic values and be a great contribution to the field of earth observation.

Bibliography

[1] V Singhroy and K Molch. "Characterizing and monitoring rockslides from SAR techniques." In: *Advances in Space Research* 33.3 (2004). Monitoring of Changes Related to Natural and Manmade Hazards Using Space Technology, pp. 290–295. ISSN: 0273-1177. DOI: `https://doi.org/10.1016/S0273-1177(03)00470-8`. URL: `https://www.sciencedirect.com/science/article/pii/S0273117703004708`.

[2] P. Matgen et al. "Towards an automated SAR-based flood monitoring system: Lessons learned from two case studies." In: *Physics and Chemistry of the Earth, Parts A/B/C* 36.7 (2011). Recent Advances in Mapping and Modelling Flood Processes in Lowland Areas, pp. 241–252. ISSN: 1474-7065. DOI: `https://doi.org/10.1016/j.pce.2010.12.009`. URL: `https://www.sciencedirect.com/science/article/pii/S1474706510002160`.

[3] R.K. Sharma et al. "SAR for disaster management." In: *IEEE Aerospace and Electronic Systems Magazine* 23.6 (2008), pp. 4–9. DOI: `10.1109/MAES.2008.4558001`.

[4] Chang an LIU et al. "Research advances of SAR remote sensing for agriculture applications: A review." In: *Journal of Integrative Agriculture* 18.3 (2019), pp. 506–525. ISSN: 2095-3119. DOI: `https://doi.org/10.1016/S2095-3119(18)62016-7`. URL: `https://www.sciencedirect.com/science/article/pii/S2095311918620167`.

[5] Miao Kang et al. "Contextual Region-Based Convolutional Neural Network with Multilayer Fusion for SAR Ship Detection." In: *Remote Sensing* 9.8 (2017). ISSN: 2072-4292. DOI: `10.3390/rs9080860`. URL: `https://www.mdpi.com/2072-4292/9/8/860`.

[6] Martin Gade and Werner Alpers. "Using ERS-2 SAR images for routine observation of marine pollution in European coastal waters." In: *Science of The Total Environment* 237-238 (1999), pp. 441–448. ISSN: 0048-9697. DOI: `https://doi.org/10.1016/S0048-9697(99)00156-4`. URL: `https://www.sciencedirect.com/science/article/pii/S0048969799001564`.

[7] Motofumi Arii, Masakazu Koiwa, and Yoshifumi Aoki. "Applicability of SAR to Marine Debris Surveillance After the Great East Japan Earthquake." In: *IEEE Journal of Selected Topics in Applied Earth Observations*

and Remote Sensing 7.5 (2014), pp. 1729–1744. DOI: `10.1109/JSTARS.2014.2308550`.

[8] Jianchao Fan et al. "Oil Spill Monitoring Based on SAR Remote Sensing Imagery." In: *Aquatic Procedia* 3 (2015). Maritime Oil Spill Response, pp. 112–118. ISSN: 2214-241X. DOI: `https://doi.org/10.1016/j.aqpro.2015.02.234`. URL: `https://www.sciencedirect.com/science/article/pii/S2214241X15002357`.

[9] Paul F Kingston. "Long-term Environmental Impact of Oil Spills." In: *Spill Science Technology Bulletin* 7.1 (2002), pp. 53–61. ISSN: 1353-2561. DOI: `https://doi.org/10.1016/S1353-2561(02)00051-8`. URL: `https://www.sciencedirect.com/science/article/pii/S1353256102000518`.

[10] Ø. Langangen et al. "The effects of oil spills on marine fish: Implications of spatial variation in natural mortality." In: *Marine Pollution Bulletin* 119.1 (2017), pp. 102–109. ISSN: 0025-326X. DOI: `https://doi.org/10.1016/j.marpolbul.2017.03.037`. URL: `https://www.sciencedirect.com/science/article/pii/S0025326X17302552`.

[11] Jones CE Brekke C. "SAR oil spill imaging, interpretation and information retrieval techniques." In: *Maritime Surveillance with Synthetic Aperture Radar* (2020), pp. 227–267. DOI: `https://doi.org/10.1049/SBRA521E_ch9`. URL: `https://hdl.handle.net/10037/20750`.

[12] Zhe Zhu et al. "Benefits of the free and open Landsat data policy." In: *Remote Sensing of Environment* 224 (2019), pp. 382–385. ISSN: 0034-4257. DOI: `https://doi.org/10.1016/j.rse.2019.02.016`. URL: `https://www.sciencedirect.com/science/article/pii/S0034425719300719`.

[13] A.H.S. Solberg et al. "Automatic detection of oil spills in ERS SAR images." In: *IEEE Transactions on Geoscience and Remote Sensing* 37.4 (1999), pp. 1916–1924. DOI: `10.1109/36.774704`.

[14] A.H. Schistad Solberg and R. Solberg. "A large-scale evaluation of features for automatic detection of oil spills in ERS SAR images." In: *IGARSS '96. 1996 International Geoscience and Remote Sensing Symposium*. Vol. 3. 1996, 1484–1486 vol.3. DOI: `10.1109/IGARSS.1996.516705`.

[15] Filippo Maria Bianchi, Martine M. Espeseth, and Njål Borch. "Large-Scale Detection and Categorization of Oil Spills from SAR Images with Deep Learning." In: *Remote Sensing* 12.14 (2020). ISSN: 2072-4292. DOI: `10.3390/rs12142260`. URL: `https://www.mdpi.com/2072-4292/12/14/2260`.

[16] F. Nirchio et al. "Automatic detection of oil spills from SAR images." In: *International Journal of Remote Sensing* 26.6 (2005), pp. 1157–1174. DOI: `10.1080/01431160512331326558`. eprint: `https://doi.org/10.1080/01431160512331326558`. URL: `https://doi.org/10.1080/01431160512331326558`.

[17] H. Espedal. "Detection of oil spill and natural film in the marine environment by spaceborne SAR." In: *IEEE 1999 International Geoscience*

and Remote Sensing Symposium. IGARSS'99 (Cat. No.99CH36293). Vol. 3. 1999, 1478–1480 vol.3. DOI: 10.1109/IGARSS.1999.771993.

[18] T. Strømgren et al. "Acute toxic effects of produced water in relation to chemical composition and dispersion." In: *Marine Environmental Research* 40.2 (1995), pp. 147–169. ISSN: 0141-1136. DOI: https://doi.org/10.1016/0141-1136(94)00143-D. URL: https://www.sciencedirect.com/science/article/pii/014111369400143D.

[19] Rodrigo N. Vasconcelos et al. "Oil Spill Detection and Mapping: A 50-Year Bibliometric Analysis." In: *Remote Sensing* 12.21 (2020). ISSN: 2072-4292. DOI: 10.3390/rs12213647. URL: https://www.mdpi.com/2072-4292/12/21/3647.

[20] Suman Singha, Domenico Velotto, and Susanne Lehner. "Near real time monitoring of platform sourced pollution using TerraSAR-X over the North Sea." In: *Marine pollution bulletin* 86.1-2 (2014), pp. 379–390.

[21] Stine Skrunes, A. Malin Johansson, and Camilla Brekke. "Synthetic Aperture Radar Remote Sensing of Operational Platform Produced Water Releases." In: *Remote Sensing* 11.23 (2019). ISSN: 2072-4292. DOI: 10.3390/rs11232882. URL: https://www.mdpi.com/2072-4292/11/23/2882.

[22] HA Espedal and OM Johannessen. "Cover: detection of oil spills near offshore installations using synthetic aperture radar (SAR)." In: (2000).

[23] Shamsudeen Temitope Yekeen, Abdul-Lateef Balogun, and Khamaruzaman B. Wan Yusof. "A novel deep learning instance segmentation model for automated marine oil spill detection." In: *ISPRS Journal of Photogrammetry and Remote Sensing* 167 (2020), pp. 190–200. ISSN: 0924-2716. DOI: https://doi.org/10.1016/j.isprsjprs.2020.07.011. URL: https://www.sciencedirect.com/science/article/pii/S0924271620301982.

[24] Charles Elachi and Jakob van Zyl. "Imaging Sensors: Synthetic-Aperture Radars." In: *Introduction to the Physics and Techniques of Remote Sensing*. John Wiley & Sons, Ltd, 2006. Chap. 6, pp. 249–291. ISBN: 9780471783398. DOI: https://doi.org/10.1002/0471783390.ch6. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/0471783390.ch6. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/0471783390.ch6.

[25] Nadav Levanon. "Radar principles." In: *New York* (1988).

[26] Merrill I Skolnik. "Introduction to radar." In: *Radar handbook* 2 (1962), p. 21.

[27] Ronald F Woodman. "Coherent radar imaging: Signal processing and statistical properties." In: *Radio Science* 32.6 (1997), pp. 2373–2391.

[28] Eugene F Knott, John F Schaeffer, and Michael T Tulley. *Radar cross section*. SciTech Publishing, 2004.

[29] Stéphane Guillaso et al. "Range resolution improvement of airborne SAR images." In: *IEEE Geoscience and Remote Sensing Letters* 3.1 (2006), pp. 135–139.

[30] ESA. *Sentinel-1*. Available at https://sentinel.esa.int/web/sentinel/missions/sentinel-1. 2014, accessed October 16, 2020. (Visited on 10/16/2020).

[31] Andrea Monti Guarnieri and Claudio Prati. "ScanSAR focusing and interferometry." In: *IEEE transactions on geoscience and remote sensing* 34.4 (1996), pp. 1029–1038.

[32] S. R. Cloude and E. Pottier. "A review of target decomposition theorems in radar polarimetry." In: *IEEE Transactions on Geoscience and Remote Sensing* 34.2 (1996), pp. 498–518. DOI: `10.1109/36.485127`.

[33] Jong-Sen Lee, M. R. Grunes, and E. Pottier. "Quantitative comparison of classification capability: fully polarimetric versus dual and single-polarization SAR." In: *IEEE Transactions on Geoscience and Remote Sensing* 39.11 (2001), pp. 2343–2351. DOI: `10.1109/36.964970`.

[34] Carlos Díaz-Avalos, Pablo Juan, and Jorge Mateu. "Similarity measures of conditional intensity functions to test separability in multidimensional point processes." In: *Stochastic Environmental Research and Risk Assessment* 27.5 (2013), pp. 1193–1205.

[35] W.M. Boerner et al. "Critical Assessment of diverse Polarimetric SAR systems - pros and cons." In: (June 2010).

[36] J. Folkesson, H. Chang, and N. Bore. "Lambert's Cosine Law and Sidescan Sonar Modeling." In: *2020 IEEE/OES Autonomous Underwater Vehicles Symposium (AUV)*. 2020, pp. 1–6. DOI: `10.1109/AUV50043.2020.9267946`.

[37] A. Cristea, J. van Houtte, and A. P. Doulgeris. "Integrating Incidence Angle Dependencies Into the Clustering-Based Segmentation of SAR Images." In: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 13 (2020), pp. 2925–2939. DOI: `10.1109/JSTARS.2020.2993067`.

[38] A.P. Doulgeris and A. Cristea. "Incorporating incidence angle variation into sar image segmentation." In: *2018 IEEE International Symposium on Geoscience and Remote Sensing*. 2018, pp. 8552–8555.

[39] George W. Swenson and A. Richard Thompson. "34 - Radio Noise and Interference." In: *Reference Data for Engineers (Ninth Edition)*. Ed. by Wendy M. Middleton and Mac E. Van Valkenburg. Ninth Edition. Woburn: Newnes, 2002, pp. 34–1–34–13. ISBN: 978-0-7506-7291-7. DOI: `https://doi.org/10.1016/B978-075067291-7/50036-4`. URL: `https://www.sciencedirect.com/science/article/pii/B9780750672917500364`.

[40] Mohammad A. Al-Ghouti et al. "Produced water characteristics, treatment and reuse: A review." In: *Journal of Water Process Engineering* 28 (2019), pp. 222 –239. ISSN: 2214-7144. DOI: `https://doi.org/10.1016/j.jwpe.2019.02.001`. URL: `http://www.sciencedirect.com/science/article/pii/S2214714418306858`.

[41] Jerry Neff, Kenneth Lee, and Elisabeth Deblois. "Produced Water: Overview of Composition, Fates, and Effects." In: July 2011, pp. 3–54. ISBN: 978-1-4614-0045-5. DOI: 10.1007/978-1-4614-0046-2_1.

[42] S. Johnsen et al. "The Environmental Impact Factor - a proposed tool for produced water impact reduction, management and regulation." In: vol. All Days. SPE International Conference and Exhibition on Health, Safety, Environment, and Sustainability. SPE-61178-MS. June 2000. DOI: 10.2118/61178-MS. eprint: https://onepetro.org/SPEHSE/proceedings-pdf/00HSE/All-00HSE/SPE-61178-MS/1902341/spe-61178-ms.pdf. URL: https://doi.org/10.2118/61178-MS.

[43] Ebenezer T. Igunnu and George Z. Chen. "Produced water treatment technologies." In: *International Journal of Low-Carbon Technologies* 9.3 (July 2012), pp. 157–177. ISSN: 1748-1317. DOI: 10.1093/ijlct/cts049. eprint: https://academic.oup.com/ijlct/article-pdf/9/3/157/9644076/cts049.pdf. URL: https://doi.org/10.1093/ijlct/cts049.

[44] Cloelle Danforth et al. "Alternative Management of Oil and Gas Produced Water Requires More Research on Its Hazards and Risks." In: *Integrated Environmental Assessment and Management* 15.5 (2019), pp. 677–682. DOI: https://doi.org/10.1002/ieam.4160. eprint: https://setac.onlinelibrary.wiley.com/doi/pdf/10.1002/ieam.4160. URL: https://setac.onlinelibrary.wiley.com/doi/abs/10.1002/ieam.4160.

[45] Petroleumstilsynet. *Aktivitetsforskriften §60*. Lov. Det Kongelige Arbeids- og Sosialdepartementet, May 2019. URL: https://www.ptil.no/regelverk/alle-forskrifter/aktivitetsforskriften/XI/60/.

[46] F. Girard-Ardhuin, G. Mercier, and R. Garello. "Oil slick detection by SAR imagery: potential and limitation." In: *Oceans 2003. Celebrating the Past ... Teaming Toward the Future (IEEE Cat. No.03CH37492)*. Vol. 1. 2003, 164–169 Vol.1. DOI: 10.1109/OCEANS.2003.178539.

[47] Y. Liu and D. K. P. Yue. "On generalized Bragg scattering of surface waves by bottom ripples." In: *Journal of Fluid Mechanics* 356 (1998), 297–326. DOI: 10.1017/S0022112097007969.

[48] Paul H LeBlond and Lawrence A Mysak. *Waves in the Ocean*. Elsevier, 1981.

[49] E.H Lucassen-Reynders, A Cagna, and J Lucassen. "Gibbs elasticity, surface dilational modulus and diffusional relaxation in nonionic surfactant monolayers." In: *Colloids and Surfaces A: Physicochemical and Engineering Aspects* 186.1 (2001), pp. 63–72. ISSN: 0927-7757. DOI: https://doi.org/10.1016/S0927-7757(01)00483-6. URL: https://www.sciencedirect.com/science/article/pii/S0927775701004836.

[50] Brent Minchew, Cathleen E. Jones, and Benjamin Holt. "Polarimetric Analysis of Backscatter From the Deepwater Horizon Oil Spill Using L-Band Synthetic Aperture Radar." In: *IEEE Transactions on Geoscience and Remote Sensing* 50.10 (2012), pp. 3812–3830. DOI: 10.1109/TGRS.2012.2185804.

[51] Roger Lang et al. “Accurate measurements of the dielectric constant of seawater at L band.” In: *Radio Science* 51.1 (2016), pp. 2–24. DOI: https://doi.org/10.1002/2015RS005776. eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2015RS005776. URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2015RS005776.

[52] Brent Minchew. “Determining the mixing of oil and sea water using polarimetric synthetic aperture radar.” In: *Geophysical Research Letters* 39.16 (). DOI: https://doi.org/10.1029/2012GL052304. eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2012GL052304. URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2012GL052304.

[53] *Bonn Agreement, Guidelines for Oil Pollution Detection: Investigation and Post Flight Analysis/Evaluation for Volume Estimation. 2017. Available online: https://www.bonnagreement.org/ (accessed on 10.06.2021).*

[54] Merv Fingas. “The challenges of remotely measuring oil slick thickness.” In: *Remote sensing* 10.2 (2018), p. 319.

[55] V. Wismann et al. “Radar signatures of marine mineral oil spills measured by an airborne multi-frequency radar.” In: *International Journal of Remote Sensing* 19.18 (1998), pp. 3607–3623. DOI: 10.1080/014311698213849. eprint: https://doi.org/10.1080/014311698213849. URL: https://doi.org/10.1080/014311698213849.

[56] M. M. Espeseth et al. “Oil-Spill-Response-Oriented Information Products Derived From a Rapid-Repeat Time Series of SAR Images.” In: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 13 (2020), pp. 3448–3461. DOI: 10.1109/JSTARS.2020.3003686.

[57] Irina Sergievskaya et al. “Damping of surface waves due to crude oil/oil emulsion films on water.” In: *Marine Pollution Bulletin* 146 (2019), pp. 206–214. ISSN: 0025-326X. DOI: https://doi.org/10.1016/j.marpolbul.2019.06.018. URL: https://www.sciencedirect.com/science/article/pii/S0025326X19304643.

[58] D. Schuler, J. Lee, and T. Ainsworth. “Polarimetric SAR Detection of Man-Made Structures Using Normalized Circular-pol Correlation Coefficients.” In: *2006 IEEE International Symposium on Geoscience and Remote Sensing*. 2006, pp. 485–488. DOI: 10.1109/IGARSS.2006.129.

[59] H. A. Espedal. “Satellite SAR oil spill detection using wind history information.” In: *International Journal of Remote Sensing* 20.1 (1999), pp. 49–65. DOI: 10.1080/014311699213596. eprint: https://doi.org/10.1080/014311699213596. URL: https://doi.org/10.1080/014311699213596.

[60] G Udny Yule. “Notes on the theory of association of attributes in statistics.” In: *Biometrika* 2.2 (1903), pp. 121–134.

[61] Martin Thoma. “A survey of semantic segmentation.” In: *arXiv preprint arXiv:1602.06541* (2016).

[62] Yi Zhao et al. “An overview of object detection and tracking.” In: *2015 IEEE International Conference on Information and Automation*. IEEE. 2015, pp. 280–286.

[63] Abdul Mueed Hafiz and Ghulam Mohiuddin Bhat. “A survey on instance segmentation: state of the art.” In: *International Journal of Multimedia Information Retrieval* (2020), pp. 1–19.

[64] Nikhil R Pal and Sankar K Pal. “A review on image segmentation techniques.” In: *Pattern Recognition* 26.9 (1993), pp. 1277 –1294. ISSN: 0031-3203. DOI: `https://doi.org/10.1016/0031-3203(93)90135-J`. URL: `http://www.sciencedirect.com/science/article/pii/003132039390135J`.

[65] Marcel Minnaert. “The reciprocity principle in lunar photometry.” In: *The Astrophysical Journal* 93 (1941), pp. 403–410.

[66] Kaiming He et al. *Mask R-CNN*. 2018. arXiv: `1703.06870 [cs.CV]`.

[67] Shaoqing Ren et al. “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks.” In: *CoRR* abs/1506.01497 (2015). arXiv: `1506.01497`. URL: `http://arxiv.org/abs/1506.01497`.

[68] Jeremiah W. Johnson. “Adapting Mask-RCNN for Automatic Nucleus Segmentation.” In: *CoRR* abs/1805.00500 (2018). arXiv: `1805.00500`. URL: `http://arxiv.org/abs/1805.00500`.

[69] Phil Ammirato and Alexander C. Berg. “A Mask-RCNN Baseline for Probabilistic Object Detection.” In: *CoRR* abs/1908.03621 (2019). arXiv: `1908.03621`. URL: `http://arxiv.org/abs/1908.03621`.

[70] Evi Kopelowitz and Guy Engelhard. *Lung Nodules Detection and Segmentation Using 3D Mask-RCNN*. 2019. arXiv: `1907.07676 [eess.IV]`.

[71] Tsung-Yi Lin et al. “Feature Pyramid Networks for Object Detection.” In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2017.

[72] Wenhao He et al. “Realtime multi-scale scene text detection with scale-based region proposal network.” In: *Pattern Recognition* 98 (2020), p. 107026. ISSN: 0031-3203. DOI: `https://doi.org/10.1016/j.patcog.2019.107026`. URL: `https://www.sciencedirect.com/science/article/pii/S0031320319303292`.

[73] Waleed Abdulla. *Mask R-CNN for object detection and instance segmentation on Keras and TensorFlow*. `https://github.com/matterport/Mask_RCNN`. 2017.

[74] Amri Emna et al. “Offshore Oil Slicks Detection From SAR Images Through The Mask-RCNN Deep Learning Model.” In: *2020 International Joint Conference on Neural Networks (IJCNN)*. 2020, pp. 1–8. DOI: `10.1109/IJCNN48605.2020.9206652`.

[75] Werner Alpers and Heinrich Hühnerfuss. “The damping of ocean waves by surface films: A new look at an old problem.” In: *Journal of Geophysical Research: Oceans* 94.C5 (1989), pp. 6251–6265.

[76] Yongsheng Wu et al. “Effects of rainfall on oil droplet size and the dispersion of spilled oil with application to Douglas Channel, British Columbia, Canada.” In: *Marine pollution bulletin* 114 (Jan. 2017), pp. 176–182. DOI: `10.1016/j.marpolbul.2016.08.067`.

[77] Tae-Ho Kim et al. “Analysis of the contribution of wind drift factor to oil slick movement under strong tidal condition: Hebei Spirit oil spill case.” In: *PloS one* 9.1 (2014), e87393.

[78] Lifan Liu, Bin Wang, and Liming Zhang. “Decomposition of mixed pixels based on bayesian self-organizing map and Gaussian mixture model.” In: *Pattern Recognition Letters* 30.9 (2009). Advanced Intelligent Computing Theory and Methodology, pp. 820–826. ISSN: 0167-8655. DOI: `https://doi.org/10.1016/j.patrec.2008.05.026`. URL: `https://www.sciencedirect.com/science/article/pii/S0167865508001876`.

[79] Xiaomeng Geng et al. “Ship Detection and Feature Visualization Analysis Based on Lightweight CNN in VH and VV Polarization Images.” In: *Remote Sensing* 13.6 (2021). ISSN: 2072-4292. DOI: `10.3390/rs13061184`. URL: `https://www.mdpi.com/2072-4292/13/6/1184`.

[80] Rafael Barea et al. “Integrating State-of-the-Art CNNs for Multi-Sensor 3D Vehicle Detection in Real Autonomous Driving Environments.” In: *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*. IEEE. 2019, pp. 1425–1431.

[81] HVLC Gamage, WOKIS Wijesinghe, and Indika Perera. “Instance-based segmentation for boundary detection of neuropathic ulcers through Mask-RCNN.” In: *International Conference on Artificial Neural Networks*. Springer. 2019, pp. 511–522.

[82] Jeremiah W Johnson. “Adapting mask-rcnn for automatic nucleus segmentation.” In: *arXiv preprint arXiv:1805.00500* (2018).

[83] Xin Wu, Shiguang Wen, and Yuan-Ai Xie. “Improvement of Mask-RCNN Object Segmentation Algorithm.” In: Aug. 2019, pp. 582–591. ISBN: 978-3-030-27525-9. DOI: `10.1007/978-3-030-27526-6_51`.

[84] Naoki Degawa, Xin Lu, and Akio Kimura. “A performance improvement of Mask R-CNN using region proposal expansion.” In: Mar. 2019. DOI: `10.1117/12.2521383`.

[85] Mate Kisantal et al. “Augmentation for small object detection.” In: *arXiv preprint arXiv:1902.07296* (2019).

[86] Cathleen E Jones et al. “Measurement and modeling of oil slick transport.” In: *Journal of Geophysical Research: Oceans* 121.10 (2016), pp. 7759–7775.

[87] Jiange Liu et al. “Automatic detection of the ice edge in SAR imagery using curvelet transform and active contour.” In: *Remote Sensing* 8.6 (2016), p. 480.

[88] Yang Liu et al. “SAR ship detection using sea-land segmentation-based convolutional neural network.” In: *2017 International Workshop on Remote Sensing with Intelligent Processing (RSIP)*. IEEE. 2017, pp. 1–4.