



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

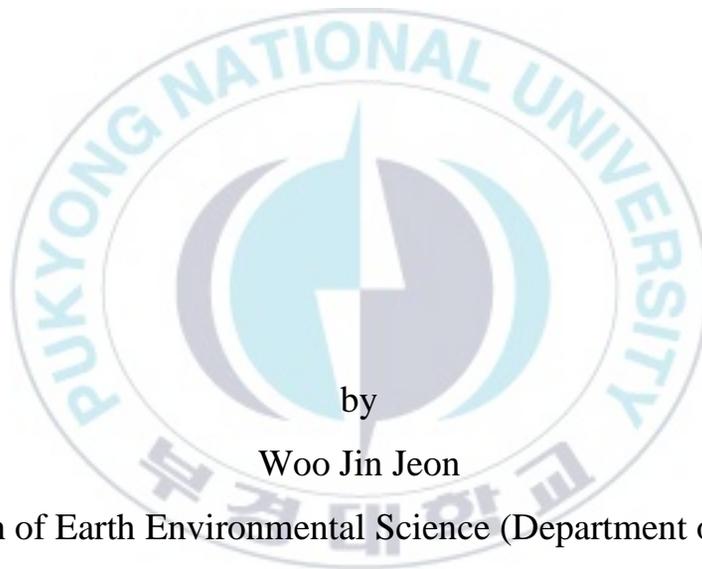
저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

Thesis for the Degree of Master of Engineering

Ship -Wake Joint Detection Using Sentinel-2A-2B Satellite Image



by

Woo Jin Jeon

Division of Earth Environmental Science (Department of Spatial

Information Engineering)

The Graduate School

Pukyong National University

February 2023

Ship -Wake Joint Detection Using Sentinel-2A-2B Satellite Image

(Sentinel-2A-2B 위성 영상을 활용한
선박 및 후류 동시 탐지 연구)

Advisor: Prof. Kyung-Soo Han

by
Woo Jin Jeon

A thesis submitted in partial fulfillment of the requirements
for the degree of
Master of Engineering
in Division of Earth Environmental Science (Department of Spatial
Information Engineering), The Graduate School,
Pukyong National University

February 2023

Ship-Wake Joint Detection Using Sentinel-2A-2B Satellite Image

A dissertation

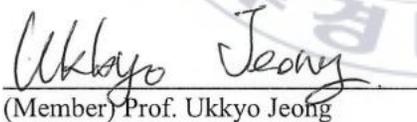
by

Woo Jin Jeon

Approved by:



(Chairman) Prof. Jin-Soo Kim



(Member) Prof. Ukkyo Jeong



(Member) Prof. Kyung-Soo Han

February 17, 2023

CONTENTS

CONTENTS	i
LIST OF FIGURES.....	ii
LIST OF TABLES.....	iv
LIST OF ACRONYMS	v
1. Introduction	1
1.1. Background	1
2. Study Area and Data.....	6
2.1. Study Area.....	6
2.2. Sentinel-2A-2B/MSI Satellite Data.....	9
2.3. Ship Reference Data	13
3. Methods for Ship Detection	15
3.1. Threshold-Based Algorithm	15
3.2. Random Forest	25
3.3. CNN-Detectron2	29
4. Results and Discussion	33
4.1. Ship Detection Results (without wake)	36
4.2. Ship Detection Results (including wake)	38
4.3. Combination of RF and CNN Results	40
5. Summary and conclusions	44
6. References	47

LIST OF FIGURES

Fig. 1. The location of ports located in Korea, and (a) is the study area of this thesis.	7
Fig. 2. Ship reference data built with Labelme and enlarged images, red border is the area of the ship reference data.	14
Fig. 3. Ship reflectance at Red, Green, Blue and NIR.	16
Fig. 4. Changes in POD and FAR according to SDI threshold.	18
Fig. 5. Ship detection results using SDI. (a) is 2021-01-19 and (b) is 2021-01-29.	20
Fig. 6. Comparison of expanded reference data and Fig. 5 ship detection results.	21
Fig. 7. Comparison of spectral characteristics of Ship and Wake in Red, Green, Blue and NIR.	23
Fig. 8. Wake Detection Index (WDI).....	24
Fig. 9. Feature importance of 14 features.	27
Fig. 10. Feature importance of the final selected features.	28
Fig. 11. Detectron2 architecture based on Mask R-CNN.	32

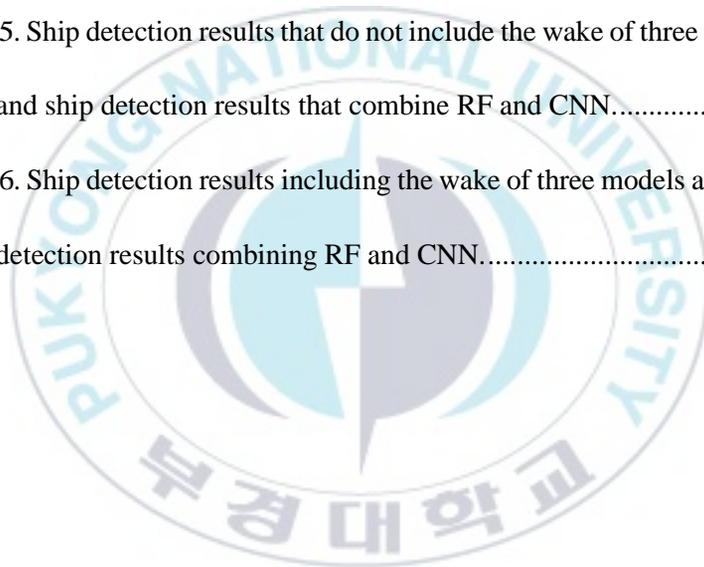
Fig. 12. Ship detection results of three models, (a) is threshold-based algorithm, (b) is Random Forest and (c) is CNN. 35

Fig. 13. Ship detection results that do not include the wake of three models. 37

Fig. 14. Ship detection results including the wake of three models..... 39

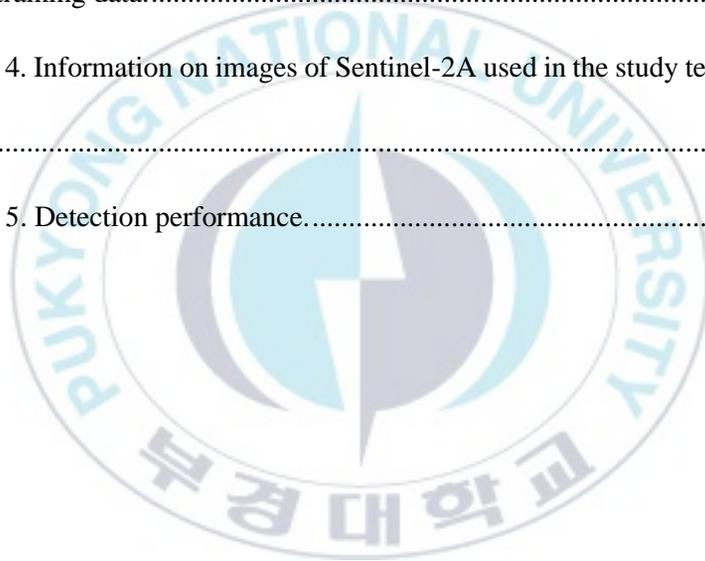
Fig. 15. Ship detection results that do not include the wake of three models and ship detection results that combine RF and CNN..... 41

Fig. 16. Ship detection results including the wake of three models and ship detection results combining RF and CNN..... 42

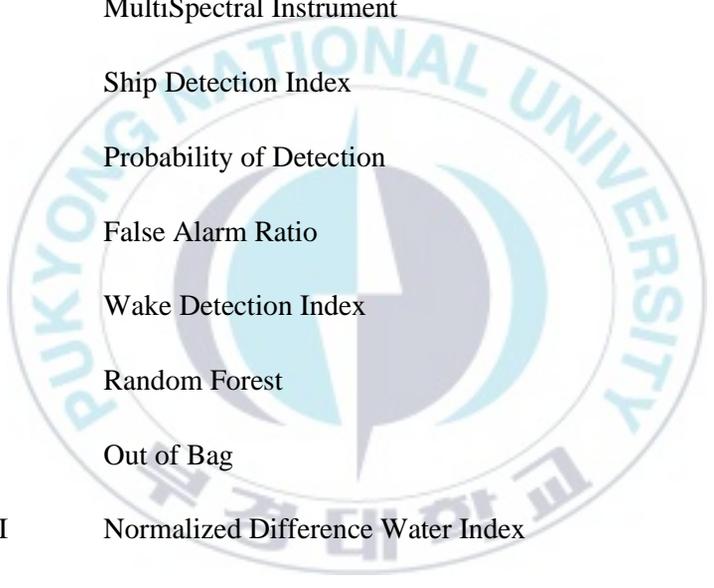


LIST OF TABLES

Table 1. Ship statistics by port for the last 2 years.....	8
Table 2. Sentinel-2A-2B/MSI channel data.....	10
Table 3. Information on images of Sentinel-2A-2B used in the study training data.....	11
Table 4. Information on images of Sentinel-2A used in the study test data.	12
Table 5. Detection performance.....	43



LIST OF ACRONYMS



AIS	Automatic Identification System
SAR	Synthetic Aperture Radar
ESA	European Space Agency
MSI	MultiSpectral Instrument
SDI	Ship Detection Index
POD	Probability of Detection
FAR	False Alarm Ratio
WDI	Wake Detection Index
RF	Random Forest
OOB	Out of Bag
NDWI	Normalized Difference Water Index
NDVI	Normalized Difference Vegetation Index
R-CNNs	Region-based Convolutional Neural Networks
YOLO	You Only Look Once
CNN	Convolutional Neural Network
RPN	Region Proposal Network
FAIR	Facebook Artificial Intelligence Research

IoU

Intersection over Union



Sentinel-2A-2B 위성 영상을 활용한 선박 및 후류 동시 탐지 연구

전 우 진

부 경 대 학 교 대 학 원 지 구 환 경 시 스템 과 학 부

공 간 정 보 시 스템 공 학 전 공

요 약

선박탐지는 해상보안, 해상교통, 어업관리, 불법조업, 국경통제 등의 분야에서 널리 사용되고 있으며, 최근 국제 해상교통량 증가로 선박사고율이 증가함에 따라 신속한 대응과 피해 최소화를 위해 선박탐지가 중요하다. 현재 다수의 세계 및 국가 규정에 따르면 특정 등급의 선박은 정기적으로 선박의 위치와 속도 등의 정보를 제공하는 자동 식별 시스템(AIS)을 장착해야 한다. 그러나 대부분의 소형 선박(300톤 미만)은 트랜스폰더를 설치할 의무가 없으며 의도적 또는 우발적으로 전송되지 않을 수 있습니다. 심지어 선박의 위치정보를 악용한 사례도 있다. 따라서 본 연구에서는 주기적으로 넓은 범위를 원격 탐지하고 소형 선박을 탐지할 수 있는 고해상도 광학위성 영상을 이용하여 선박 탐지를 수행하였다.

광학영상은 최근 극궤도, 정지궤도, 초소형 군집위성 등의 증가로 인해 이용 가능한 데이터가 다량 축적되고 있으며, 공간해상도도 향상되고 있어 다수의 고해상도 광학위성 영상을 함께 사용할 경우 검출영역을 확대하고 관측빈도를 높일 수 있다. 그러나 광학영상은 구름과 후류같이 배와 유사한 밝기를 나타내는 요인으로 인해 거짓 경보를 유발할 수 있으므로 선박 탐지의 정확도를 향상시키기 위해 이러한 요인들을 제거하는 것이 중요하다. 본 연구에서는 해상도의 한계로 인해 탐지가 어려운 소형선박의 존재나 선박의 방향과 속도를 추정하는 데 도움이 되는 후류를 제거하여 거짓 경보를 줄이고 선박 탐지의 정확도를 향상시켰다. 선박 탐지 방법으로는 가장 많이 사용되는 물체 탐지 방법인 Threshold 기반 알고리즘과 최근 물체 탐지 분야에서 널리 사용되고 있는 머

신러닝 기반의 Random Forest, CNN 기법을 이용하여 선박 탐지를 수행하였으며, 모델별 선박 탐지 결과를 비교 분석하였다. Threshold 기반 알고리즘은 선행연구의 선박탐지 지수 (Ship Detection Index, SDI)를 이용하여 본 연구영역에 적합한 경계값을 선정하고 선박 탐지를 수행하였다. SDI로 미탐지된 후류를 제거하기 위해 후류와 선박의 분광 특성을 분석하여 후류탐지 지수 (Wake Detection Index, WDI)를 개발하였고, WDI를 이용하여 후류를 제거하여 Threshold 기반 알고리즘의 선박 탐지 결과의 정확도를 향상시켰으며, RF는 선박과 후류를 동시에 감지하기 위해 밴드별 반사도, 밴드비, SDI, WDI 및 NDWI 등 총 7개의 feature를 활용하여 모델을 학습하였다. 마지막으로 CNN의 경우 Mask R-CNN 기반 Detectron2를 사용하여 선박과 후류의 이미지를 학습하여 모델을 개발하고 선박과 후류를 동시에 탐지하였다. CNN의 경우 선박이 소형으로 감지되는 경향이 있었고 Threshold 와 RF는 선박이 끊어지는 현상이 발생하였다. 따라서 본 연구에서는 RF와 CNN의 결과를 결합하여 선박끊김현상과 과소탐지현상을 개선하였다. 본 연구의 선박 탐지 결과는 높은 정확도를 유지하면서 각 모델의 한계를 개선하였다는 점에서 의의가 있다. 또한 향후 공간해상도가 향상된 위성영상을 활용할 경우 보다 정확도가 높은 선박 및 후류 동시탐지가 가능할 것으로 기대된다.

1. Introduction

1.1. Background

Ship detection is widely used in areas such as maritime security, maritime transportation, fisheries management, illegal fishing, and border control (Zou and Shi, 2016; Liu et al., 2017; Heiselberg, 2016; Li et al., 2018). In addition, ship detection is important to respond quickly and minimize damage as the rate of ship accidents increases due to the recent increase in international sea traffic.

A number of global and national regulations require ships of a particular class to be equipped with a shipborne transponder that transmit the ship's identity and location at particular repeat intervals (Kanjir et al., 2018). One of the most common tracking systems is the Automatic Identification System (AIS), designed to automatically provide location information to other vessels and offshore authorities (Kanjir et al., 2018). These systems help a lot in tracking ships, but most small (less than 300 tons) ships do not need to carry AIS. In addition, location information may not be transmitted intentionally or accidentally, and it has been observed that illegally operating vessels identify their locations and change ship orientation or even steal them (Heiselberg, 2016). Therefore, we cannot

rely entirely on systems such as AIS. This means that there is a need for a non-cooperative detection system, such as satellites.

Satellite-based sensors have the advantages of remote detection, global reach, regular updates, and high data collection volumes (Kanjir et al., 2018). Therefore, the use of satellite images is the most economical and essential tool for detecting ships in the ocean (Kanjir et al., 2018). Currently, images of optical and reflected infrared, hyperspectral, thermal infrared, and radar are used a lot. Synthetic Aperture Radar (SAR) images are widely used in ship detection because they are little affected by weather and time (Zou and Shi, 2016; Eldhuset, 1996; Dragosevic and Vachon, 2008; Li and Chong, 2008). However, in the case of SAR images, the number of satellites on board is small, so there is a limit to the area that can be covered at the same time (Kanjir et al., 2018). On the other hand, optical sensors have recently increased polar orbits, geostationary orbits, and clustered microsattellites, making it possible to observe the global at the same time and the number of data generated is increasing rapidly, interest in the potential of optical images for ocean observation is increasing exponentially (Kanjir et al., 2018). Optical images can also provide valuable information for accurate ship identification and feature extraction (Liu et al., 2013) and are relatively consistent. Add, these images are affordable and simple, enabling classification and broad

application areas (Lan and Wan, 2009). In addition, the spatial resolution of optical images is improved, so when multiple high-resolution optical satellite images are used together, the detection area can be expanded and the frequency of observation can be increased.

Ship detection in optical images can simply be considered to detect bright spots against a dark background. But the reality is that the ship may be darker than the surface of the surrounding sea, or there may be many other bright objects in the image that can be mistakenly detected as a ship. The main factors that cause false alarms are clouds, waves, and ship wakes (Yang et al., 2013). Elimination of false alarm factors is a key issue in ship detection in optical images. For example, Yang et al. (2014) compared ship detection results on a quiet-textured sea surface with a wavy sea surface and showed a marked decrease in accuracy on a noisy sea surface. Therefore, the importance of wake detection is emphasized in many studies to improve the accuracy of ship detection (Yang et al., 2011). In addition, wake detection can be used to estimate ship direction and speed and can help indicate the presence of small ships that are not identified in the image (Bouma et al., 2013; Buck et al., 2007). In other words, detecting the exact size of a ship in ship detection affects the maritime security aspect and the extent of damage to the ship's accident, so it is important to remove the obstructive wake.

The most commonly used ship detection technique distinguishes images according to the value of each pixel being higher or lower than the threshold. These methods perform relatively good performance on smooth sea surface or high contrast between ship targets and sea background, but a high increase in false alarm may occur if there is too much clutter on the image. Some studies propose methods based on shape and texture features, utilizing different characteristics of ships and sea, and typically include spectral information. Threshold-based algorithms provide relatively high detection accuracies, although false-alarm candidates (wake, clutter) still exist. The ship detection method that has been widely performed recently is a machine learning-based approach. Machine learning-based approaches are easy to learn and handle big data and yield results at very high speeds. Machine learning-based approaches are about how to organize computer programs that automatically improve with experience (Mitchell, 1997). Here, it is not necessary to explicitly define object features, but instead, the image data is used as direct input data to artificial neural network. This can be very powerful, but requires a large training set, and if the implementation is not done carefully, new objects that do not reproduce the training set may be misclassified (Kanjir et al., 2018). Therefore, this study aims to perform ship detection in optical images using Threshold-based algorithms, machine learning-based RF and CNN

models, and to compare and analyze the performance and ship detection results of the three models.



2. Study Area and Data

2.1. Study Area

The study area is 35.03806°N - 35.07056°N and 129.0278°E - 129.0878°E , including waters near Yeongdo, Busan. Korea has a high rate of ship accidents due to its complicated coastline (Song et al., 2013). According to the National Port Entry and Exit Ship Statistics (Table 1) provided by the Ministry of Maritime Affairs and Fisheries, Busan has the largest number of ships coming in and out with 292,873 ships over the past two years. Fig. 1 shows the location of ports across the country. And Yeongdo is an area with a lot of ship flow as many docks and terminals are located, such as Busan Port Coastal Passenger Terminal, Nambu Out Port, Busan International Cruise Terminal, Hari Port, and Buk Port. Therefore, we designated the Yeongdo area of Busan as a study area and conducted ship detection.

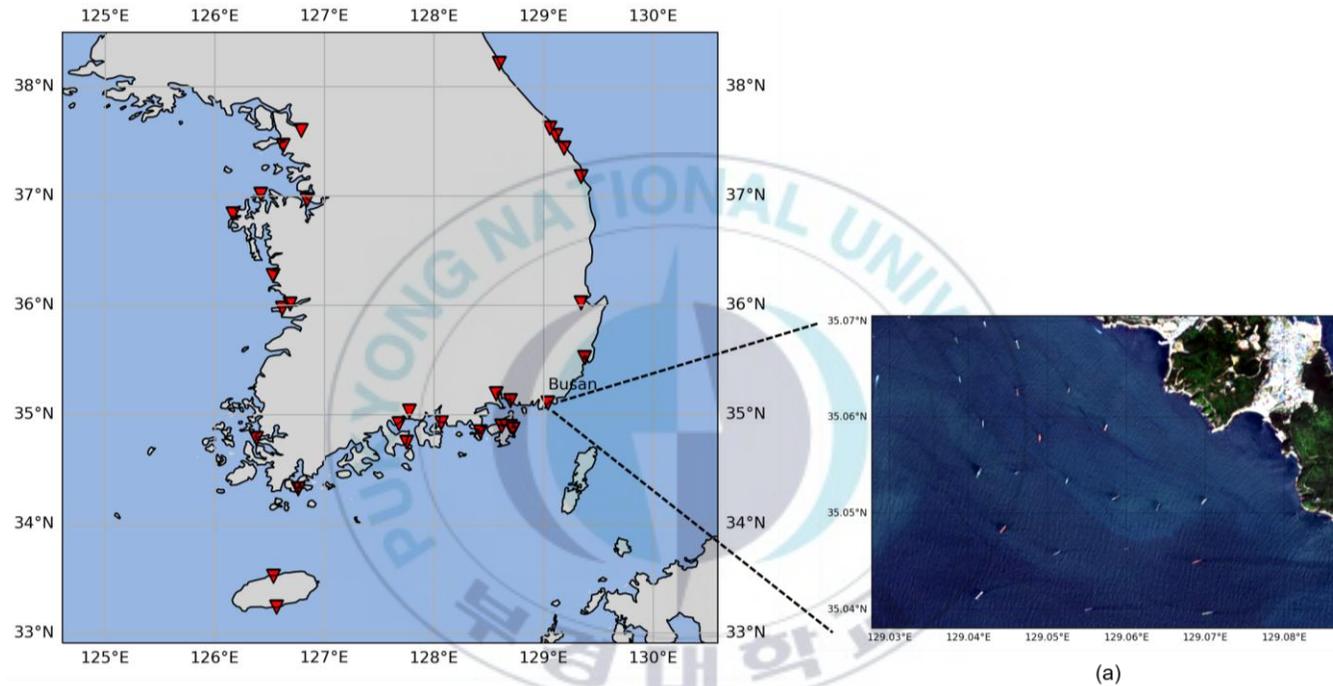


Fig. 1. The location of ports located in Korea, and (a) is the study area of this thesis.

Table 1. Ship statistics by port for the last 2 years.

Busan	Incheon	Pyeongtaek. Dangjin	Gyeongin	Donghae	Samcheo k	Sokcho	Okpye	Hosan	Daesan
292,873	101,925	57,940	1,596	25,810	7,472	1,697	8,101	1,416	45,870
Boryung	Taeon	Gunsan	Janghang	Mokpo	Wando	Yeosu	Gwngyang	Pohang	Masan
2,970	2,029	23,572	3,554	50,339	6,286	57,032	155,605	40,867	35,976
Samcheon po	Okpo	Jangseungpo	Jinhae	Tongyeong	Gohyeon	Hadong	Ulsan	Jeju	Seogwipo
5,105	13,053	93	10,817	7,368	11,069	1,430	159,187	38,062	18,458

2.2. Sentinel-2A-2B/MSI Satellite Data

The Sentinel-2 satellite was developed by the European Space Agency (ESA) for Earth observation and is currently operated by two satellites, Sentinel-2A and Sentinel-2B. Sentinel-2A and Sentinel-2B were launched on 23 June 2015 and 7 March 2017, respectively. Sentinel-2A-2B revisit period is 5 days. Table 2 shows the specification of Sentinel-2A-2B Multispectral Instrument (MSI).

Information on Sentinel-2A-2B/MSI images (Scene 1-20) for the research area shown in Fig. 1(a) is shown in Tables 3 and 4, and this information describes the date, time. The scenes in Table 3 were used as training data for the three ship detection models, and the scenes in Table 4 were used as test data for the accuracy evaluation of the prediction results of the three ship detection models. Sentinel-2A-2B/MSI images used in this study are Level 2 data, and these images are used in the study of Blue (Band2), Green (Band3), Red (Band4), and NIR (Band8) bands with cloudless spatial resolution of 10 m.

Table 2. Sentinel-2A-2B/MSI channel data.

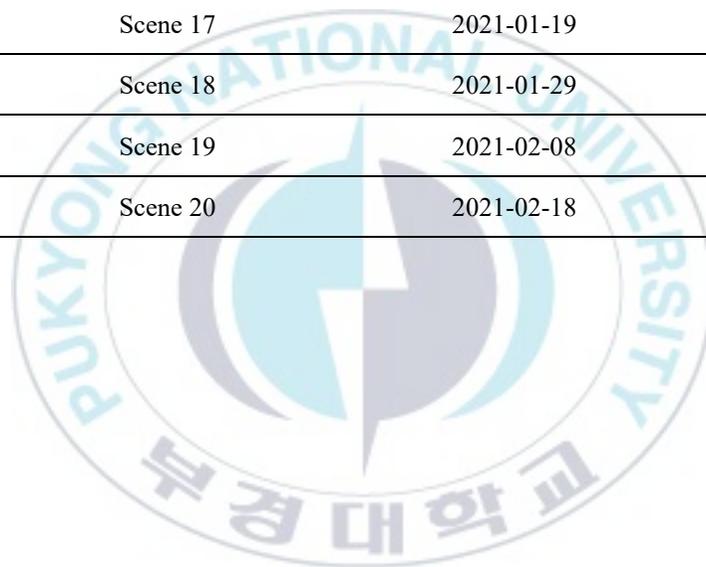
Sentinel-2 bands	Sentinel-2A		Sentinel-2B		Spatial resolution (m)
	Central wavelength (nm)	Band width (nm)	Central wavelength (nm)	Band width (nm)	
Band 1	442.7	21	442.2	21	60
Band 2	492.4	66	492.1	66	10
Band 3	559.8	36	559.0	36	10
Band 4	664.6	31	664.9	31	10
Band 5	704.1	15	703.8	16	20
Band 6	740.5	15	739.1	15	20
Band 7	782.8	20	779.7	20	20
Band 8	835.8	106	832.9	106	10
Band 8A	864.7	21	864.0	22	20
Band 9	945.1	20	943.2	21	60
Band 10	1373.5	31	1376.9	30	60
Band 11	1613.7	91	1610.4	94	20
Band 12	2202.4	175	2185.7	185	20

Table 3. Information on images of Sentinel-2A-2B used in the study training data.

Satellite	Scene number	Date	Time (UTC)
Sentinel-2A	Scene 4	2019-04-20	02:07:01
	Scene 8	2020-01-05	02:10:51
	Scene 9	2020-01-15	02:10:31
	Scene 10	2020-02-04	02:09:01
	Scene 11	2020-02-14	02:08:01
	Scene 12	2020-05-04	02:07:01
	Scene 15	2020-11-30	02:10:41
	Scene 16	2021-01-09	02:10:51
Sentinel-2B	Scene 1	2019-02-04	02:08:59
	Scene 2	2019-03-16	02:06:49
	Scene 3	2019-04-15	02:06:59
	Scene 5	2019-11-01	02:08:29
	Scene 6	2019-11-21	02:09:59
	Scene 7	2019-12-31	02:10:59
	Scene 13	2020-10-06	02:06:59
	Scene 14	2020-11-25	02:10:19

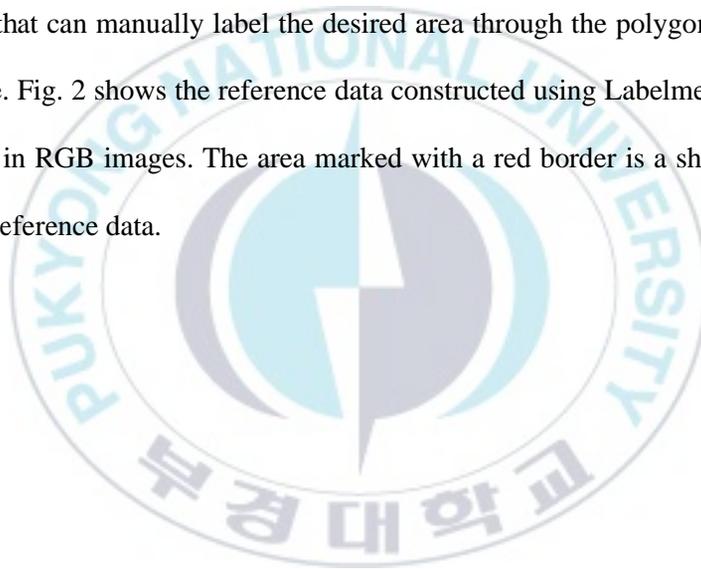
Table 4. Information on images of Sentinel-2A used in the study test data.

Satellite	Scene number	Date	Time (UTC)
Sentinel-2A	Scene 17	2021-01-19	02:10:21
	Scene 18	2021-01-29	02:09:31
	Scene 19	2021-02-08	02:08:51
	Scene 20	2021-02-18	02:07:31



2.3. Ship Reference Data

In this paper, since the Instance Segmentation method will be used for ship detection, annotation work is required for the learning dataset. Therefore, Labelme, one of the Annotation tools used for data labeling, was used to build ship reference data. Labelme is an Open Annotation Tool that can manually label the desired area through the polygon in the image. Fig. 2 shows the reference data constructed using Labelme in this study in RGB images. The area marked with a red border is a ship built with reference data.



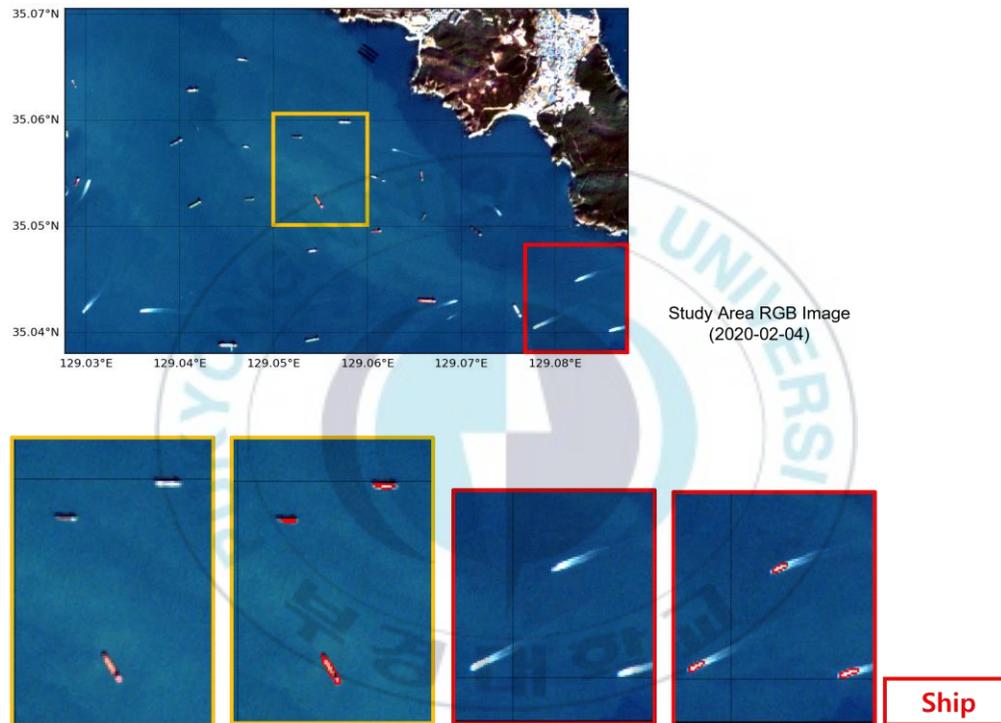


Fig. 2. Ship reference data built with Labelme and enlarged images, red border is the area of the ship reference data.

3. Methods for Ship Detection

3.1. Threshold-Based Algorithm

Ship detection based on the threshold is performed by utilizing the difference in reflectivity between the ship and the background (surrounding sea). Fig. 3 shows that ships exhibit higher reflectivity than the sea in all channels Red (Band 4), Green (Band 3), Blue (Band 2), and NIR (Band 8). The threshold - based algorithm in previous study proposed the Ship Detection Index (SDI, (Park *et al.*, 2018)) as a combination of Red and NIR, using the characteristic that the difference between ship and sea reflectance values in Red and NIR channels is relatively greater than that of Green and Blue channels.

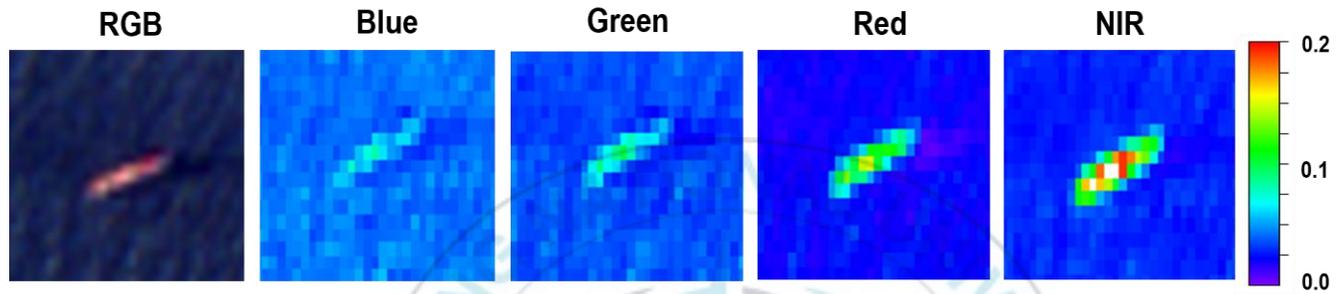


Fig. 3. Ship reflectance at Red, Green, Blue and NIR.



$$SDI = \frac{Ref_{0.66} - Ref_{0.66_min}}{Ref_{0.66_max} - Ref_{0.66_min}} \times \frac{Ref_{0.94} - Ref_{0.94_min}}{Ref_{0.94_max} - Ref_{0.94_min}} \quad (1)$$

Equation 1 is the threshold-based Ship Detection Index (SDI, (Park *et al.*, 2018)). $Ref_{0.66}$ and $Ref_{0.94}$ of Equation 1 represent the Red reflectance value and the NIR reflectance value of the pixel to be distinguished, respectively. $Ref_{0.66_min}$ and $Ref_{0.94_min}$ represent the minimum reflectance values of Red and NIR in each scene, while $Ref_{0.66_max}$ and $Ref_{0.94_max}$ represent the maximum reflectance values of Red and NIR in each scene. In this study, SDI was calculated using $Ref_{0.66_min}$ and $Ref_{0.94_min}$ as the bottom 5% values for each scene and $Ref_{0.66_max}$ and $Ref_{0.94_max}$ as the top 95% values for each scene to remove outliers and consider seasonal spectral properties during the study period. To select the optimized threshold value for this study area, we empirically modified the threshold value using the evaluation indices Probability of Detection (POD) and False Alarm Ratio (FAR). POD means that the closer to 1, the higher the accuracy, and FAR means that the closer to 0, the higher the accuracy. Fig. 4 shows the change in POD and FAR values by changing the SDI threshold value from 0.08 to 0.129

in units of 0.001. Overall, as the threshold increases, the accuracy decreases. Based on the SDI threshold of 0.0875, the slope of the POD becomes relatively gentle, and the slope of the FAR drops sharply. Therefore, in this study, the threshold of the SDI was empirically designated as 0.0875. In other words, if the SDI value is 0.0875 or more, it is classified as a ship, and if it is less than 0.0875, it is classified as a sea.

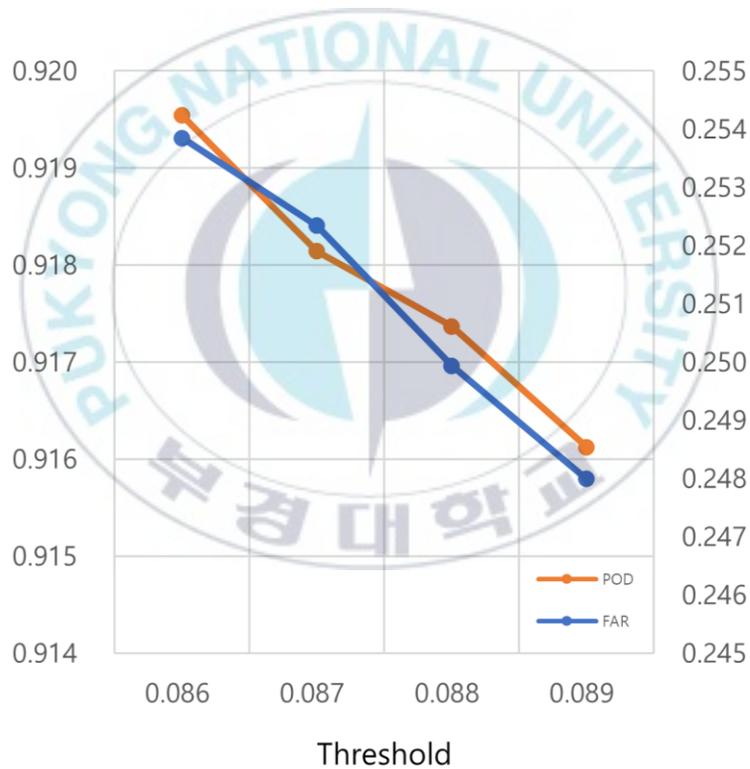
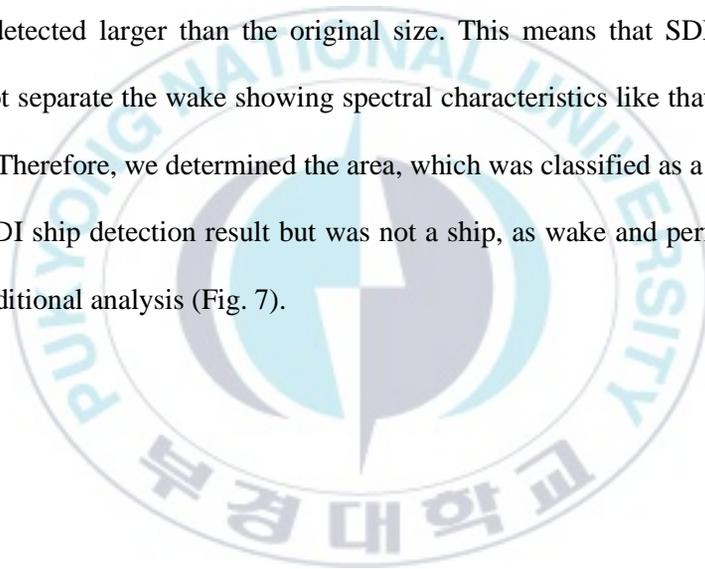


Fig. 4. Changes in POD and FAR according to SDI threshold.

Fig. 5 is the result of ship detection applying SDI to the research area, and Fig. 6, the reference data and ship detection results are shown together by expanding the ship. Through Fig. 5, it can be confirmed that ship detection using SDI is well detected in the presence or absence of a ship. But in the enlarged ship detection result of Fig. 6, it was found that the wake was detected as ship compared to the reference data, and the ship was detected larger than the original size. This means that SDI alone cannot separate the wake showing spectral characteristics like that of the ship. Therefore, we determined the area, which was classified as a ship in the SDI ship detection result but was not a ship, as wake and performed an additional analysis (Fig. 7).



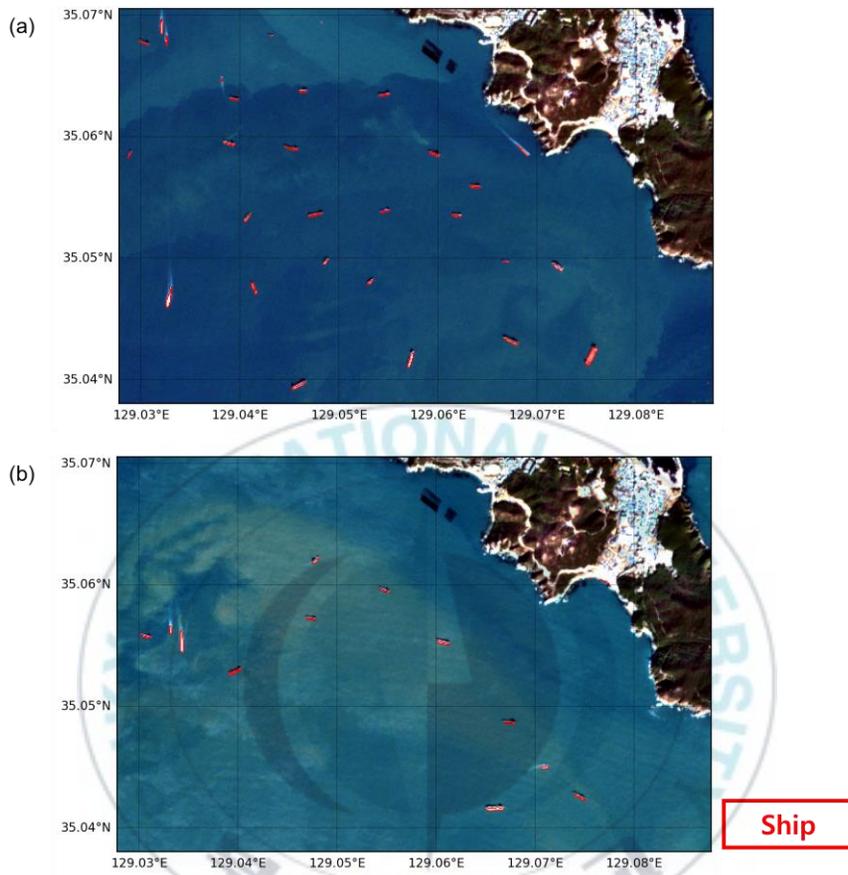


Fig. 5. Ship detection results using SDI. (a) is 2021-01-19 and (b) is 2021-01-29.

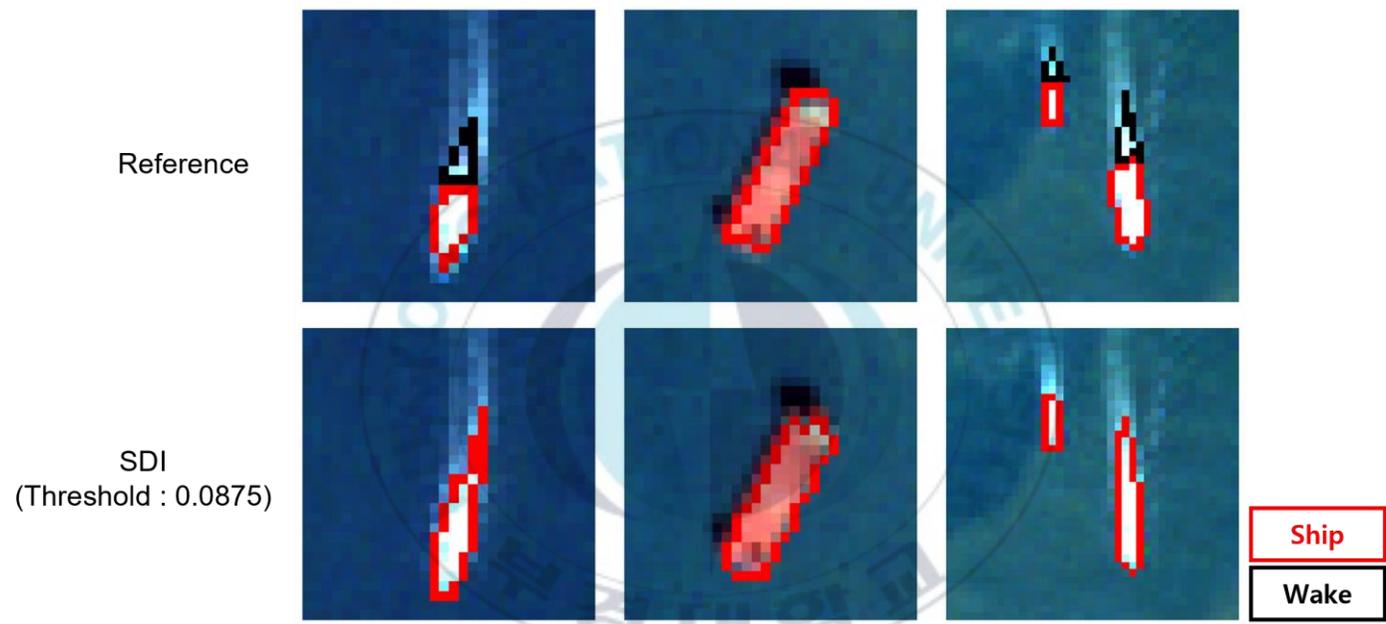


Fig. 6. Comparison of expanded reference data and Fig. 5 ship detection results.

In SDI, ship and wake show similar values and are not distinguished from each other. In addition, in Fig. 7, the wake has similar reflectance values of the ship and the Blue channel, but there is a large difference in reflectance values in the Red and NIR channels. In particular, in the case of ship, NIR channel reflectance has a greater value than Red channel reflectance, Conversely, in the case of wake, the red reflectance tends to be greater than the NIR reflectance. Therefore, based on the spectral characteristics of the wake, we developed the Wake Detection Index (WDI, Equation 2) with a combination of SDI, Blue, Red, and NIR and utilized it for the classification of the wake and the ship (Fig. 8).

$$WDI = \left(\frac{Ref_{0.46} - Ref_{0.84}}{SDI + Ref_{0.84}} \right) + \left(\frac{Ref_{0.66} - Ref_{0.84}}{SDI} \right) - 0.3 \quad (2)$$

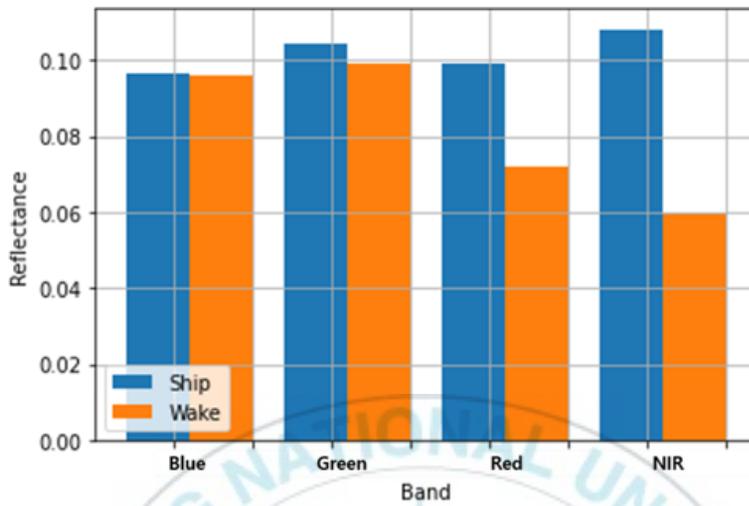


Fig. 7. Comparison of spectral characteristics of Ship and Wake in Red, Green, Blue and NIR.

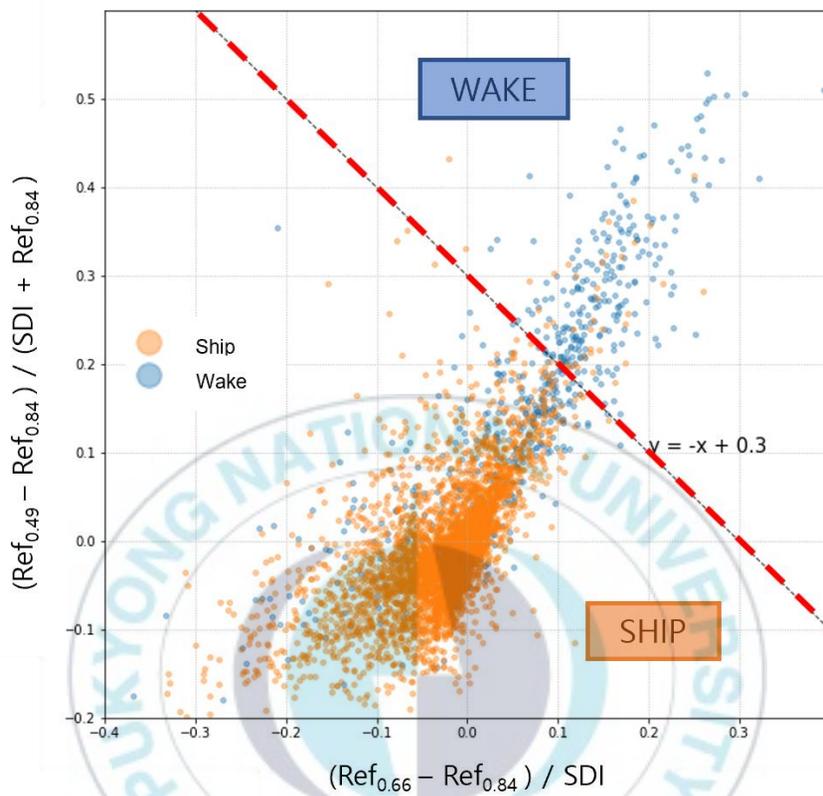


Fig. 8. Wake Detection Index (WDI)

3.2. Random Forest

Random Forest (RF) is a representative ensemble classifier based on several decision trees trained with randomly selected data subsets and feature sets (Breiman, 2001). The RF makes two random choices when generating many decision trees. First, sample a subset of training sets randomly from the training dataset and generate a decision tree from each dataset. For example, assuming that there are 1000 data in the training set, only 100 data can be arbitrarily selected to create a tree when each tree is created. That is, all trees are formed based on different data, but they are all subsets of training dataset. Second, when forming RF, it does not only change the dataset, but also changes the feature selection. When selecting a feature, a subset of existing features is used. In general, when M features exist, the number of randomly selected features utilizes the square root of M . What is important at this time is that duplication is allowed when randomly selecting data. This method is called bagging (Breiman, 1996). As a result, the problem of overfitting occurring in one decision tree model can be solved because the variance is reduced while maintaining the bias of each tree (Breiman, 2001). The final classification decision is determined by the class with the most votes among the classes calculated by all the trees generated. To estimate how well the model performed, the

RF uses approximately $2/3$ of the samples (referred to as in-bag samples) for tree learning and the remaining $1/3$ (referred to as out-of-bag) for internal cross-validation (Breiman, 2001). This estimate is called OOB (Out-of-Bag) score, and the closer the OOB score is to 1, the better the model is learned.

In this study, Sentinel-2A-2B images (Scene 1–16, 8423 ship pixels, 639 wake pixels and 2612224 sea pixels) of Table 3 were used. If the data is used as it is, learning biased toward sea characteristics can be performed due to the imbalance of the data, resulting in an overfitting model for the sea. Therefore, we adjusted the data imbalance by increasing the number of wake pixels and reducing the number of sea pixels based on the number of ship pixels. Finally, 8670 ship pixels, 7469 wake pixels and 18698 sea pixels were used as RF input data sets. First, a total of 14 input features were selected. Band reflectance and band ratio to reflect spectral characteristics, SDI that distinguishes the sea from the ship well by reflecting the spectral characteristics of the ship, and WDI that separates the wake, that is detected as the ship and generates false alarms, from the ship. Lastly, since the ship exhibits spectroscopic characteristics similar to that of the land, the Normalized Difference Water Index (NDWI), which is mainly used to analyze the water body, was used and NDVI (Normalized Difference Vegetation Index) was used to perform double

check on land. We simply made feature selection based on the relative feature importance provided by the RF. Fig. 9 shows the feature importance of 14 features. Finally, 7 parameters were selected (Fig. 10) and this was used for RF development, which simultaneously detects ships and wake and effectively removes false alarms. The OOB score of the final RF model was 0.985.

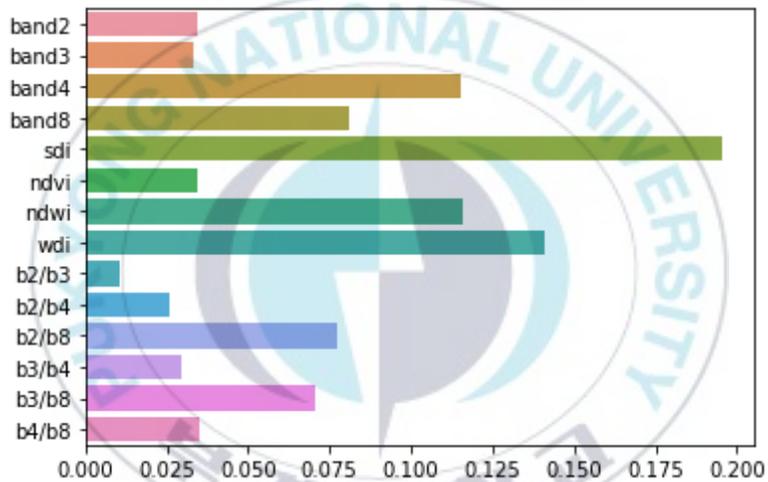


Fig. 9. Feature importance of 14 features.

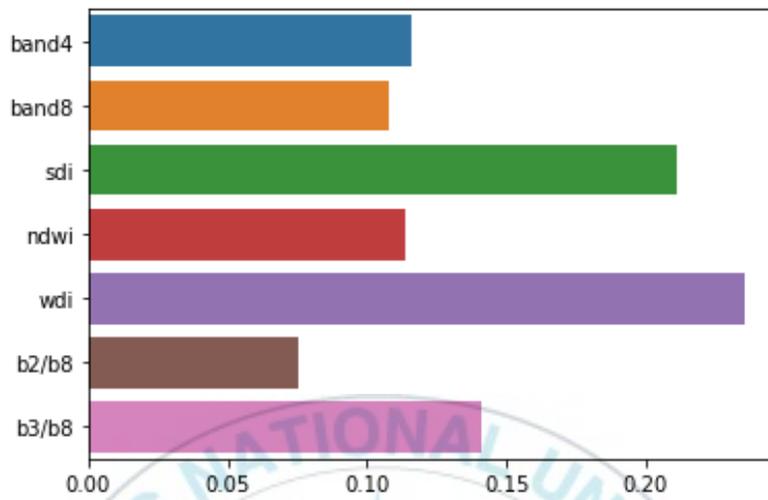
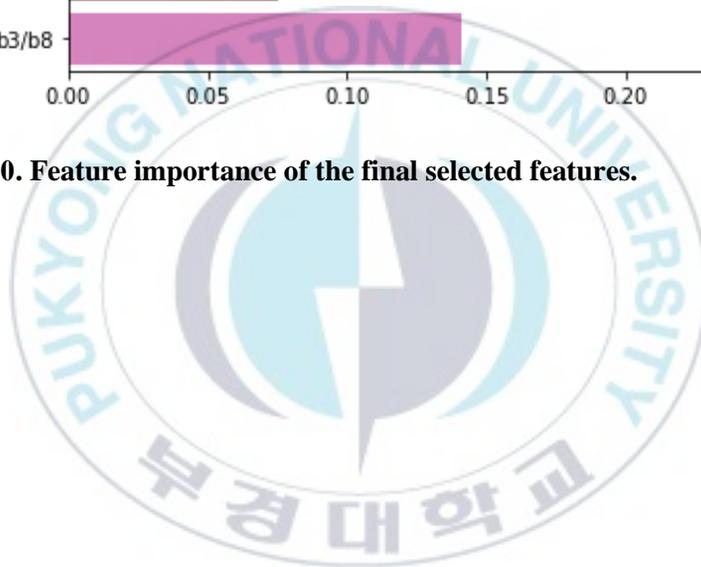


Fig. 10. Feature importance of the final selected features.

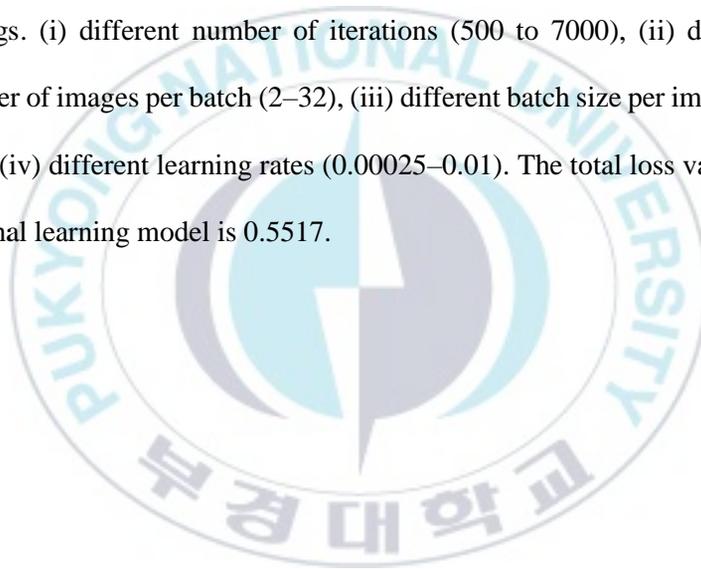


3.3. CNN-Detectron2

As a state-of-the-art method for detecting objects, models using deep neural networks are being used and performing well. Recent technologies generally relate to two models: Region-based Convolutional Neural Networks (R-CNNs) and You Only Look Once (YOLO). The R-CNNs (Ross Girshick et al., 2014) approach utilizes three modules. The first is a region proposal module that creates bounding boxes using computer vision techniques. The second is a feature extraction module. The feature extraction module uses a Convolutional Neural Network (CNN) to extract features from the candidate region. Finally, the third module is a classifier that predicts a class of proposed candidates using the extracted features. R-CNNs take a long time to train because training takes place in several stages. Therefore, Girshick proposes another model called Fast R-CNN (R. Girshick., 2015) to solve these problems. Fast R-CNNs are trained with a single model rather than three individual modules. This architecture input an image and proposes a candidate region, and then extracts features from the candidate via a popular pre-trained image classification model (e.g., ResNet (K He et al., 2016), VGG-16 (K Simonyan et al., 2014)). The biggest difference from R-CNNs is that each proposal does not go through CNN but performs object detection in the output feature map step after

going through CNN once for the entire image. Fast R-CNN improves training and prediction time, but still requires local suggestions as input. In other words, regional proposals for each image still need to be made separately. Thus, Ren et al., proposes Faster R-CNN (S Ren et al., 2015) to address this problem. A major improvement is the ability to incorporate region proposals as part of the final model using the Region Proposal Network (RPN). In other words, there are two smaller networks in this architecture. The first is RPN and the second is Fast R-CNN. These two subnetworks are trained simultaneously on two different tasks: 1) local proposal and 2) bounding box classification and regression. These strategies help improve training and object detection time and accuracy. Another famous object detection product is YOLO. Depending on the YOLO version, there may be differences in terms of the architecture and technology used. The advantage of this method is that training and prediction are faster. However, this model is slightly less accurate than Faster R-CNN (V Pham et al., 2020). Therefore, this work explores the Faster R-CNN approach to perform ship detection. Instead of developing the Faster R-CNN model from scratch, we use Detectron2 to shorten the development cycle. Detectron2 is a training/inference platform for Pytorch-based object detection and semantic segmentation created by Facebook Artificial Intelligence Research (FAIR). Detectron2 has a

structure that adds a classification branch that predicts the class of an object and a mask branch that predicts the segmentation mask parallel to the bbox regression branch that performs the bbox regression for RoI obtained from the RPN of the Faster R-CNN. The Detectron2 architecture is shown in Fig. 11. The input data is in COCO JSON format. We tested and evaluated all the proposed related algorithms with the following settings. (i) different number of iterations (500 to 7000), (ii) different number of images per batch (2–32), (iii) different batch size per image (8–512), (iv) different learning rates (0.00025–0.01). The total loss value for the final learning model is 0.5517.



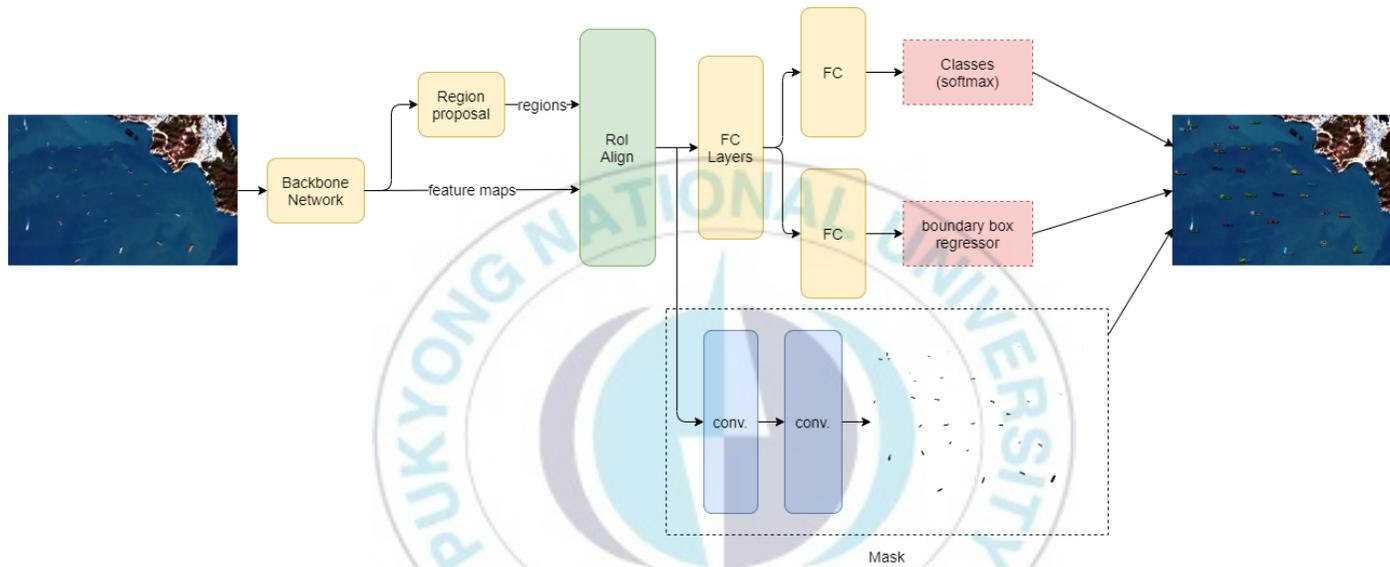


Fig. 11. Detectron2 architecture based on Mask R-CNN.

4. Results and Discussion

(a), (b), and (c) of Fig. 12 are the ship detection results of the Threshold-based algorithm, Random Forest, and CNN models, respectively. The part marked with a red border represents the edge of the pixel detected by the model as a ship. All three models were well detected without missing ships.

To analyze the results of ship detection by model in more detail, we analyzed them by dividing them into two groups: ships that do not include wake and ships that include wake. In addition, by expanding the ship, it was compared and analyzed how the area where the ship and wake were detected by model was different. In addition, in this study, Precision, Recall, F1-score, and Intersection over Union (IoU) were used to perform Evaluation for each model.

$$Precision = \frac{TruePositives}{TruePositives + FalsePositives} \quad (3)$$

Precision is the ratio of what the model classifies as Ture to what is actually Ture.

$$Recall = \frac{TurePositives}{TruePositives + FalseNegatives} \quad (4)$$

Recall is the ratio of what the model predicts to be true among what is actually True. Therefore, it is possible to determine how similar the reference data and the result predicted by the model are through the Recall.

$$F1_score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (5)$$

F1_score is the harmonic average of Precision and Recall. F1_score means that when the data label is an unbalanced structure, the performance of the model can be accurately evaluated, and the closer it is to 1, the higher the accuracy, and the closer it is to 0, the lower the accuracy.

$$IoU = \frac{Overlapping\ Region}{Combined\ Region} = \frac{Reference\ and\ Prediction}{Reference\ or\ Prediction} \quad (6)$$

IoU is an index that measures the similarity between the reference data and the result predicted by the model and has a value between 0 and 1. It is evaluated that the wider the overlapping area, the better the prediction.

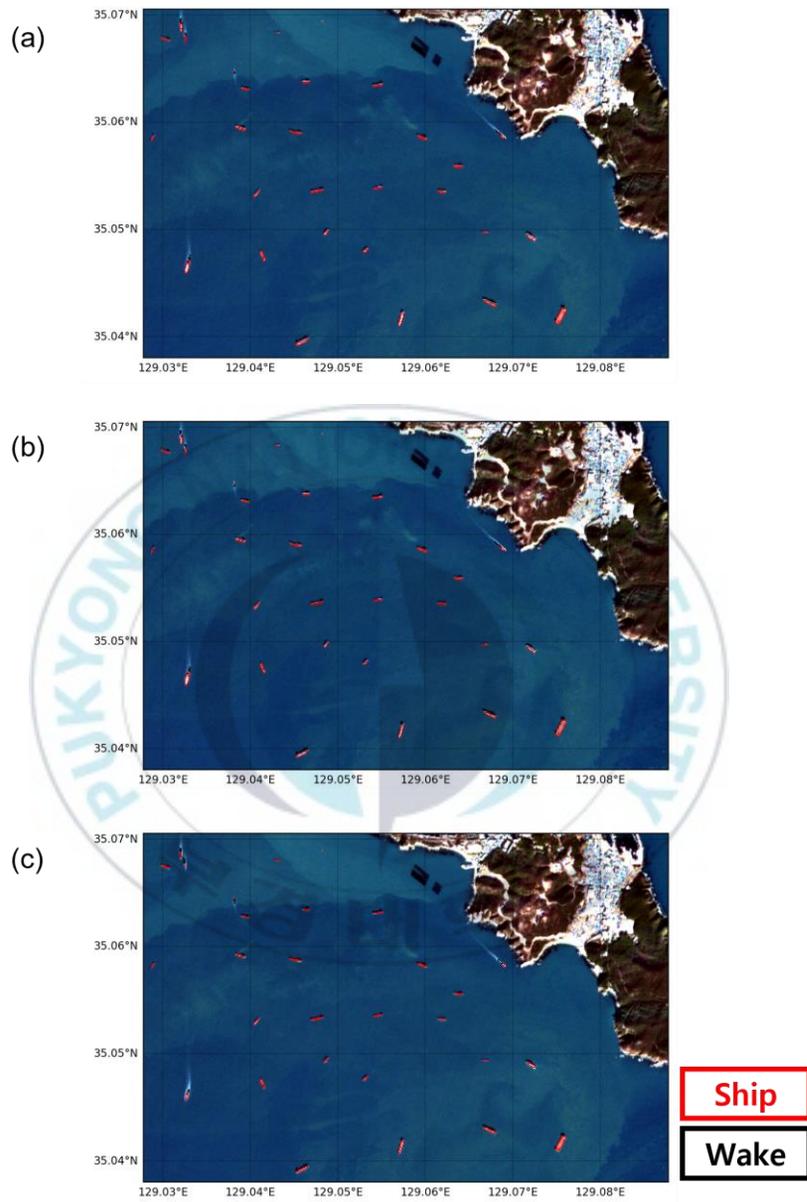


Fig. 12. Ship detection results of three models, (a) is threshold-based algorithm, (b) is Random Forest and (c) is CNN.

4.1. Ship Detection Results (without wake)

Fig. 13 is a qualitative comparison of reference data and ship detection results by model for ships that do not include a wake. The shape and area of the ship were detected differently for each model. Because Threshold-based algorithm and RF perform detection based on pixel-specific spectral characteristics, as in (i) and (n) of Fig. 13, a phenomenon in which one ship is cut off may occur. On the other hand, CNN uses Anchor box, a bounding box with various aspect ratios, by introducing RPN in the process of performing candidate region extraction tasks, so it extracts region proposals more precisely. Therefore, there was no phenomenon in which one ship was cut off, such as the threshold-based algorithm and RF. However, compared to the reference data, the size of the ship tended to be detected as small.

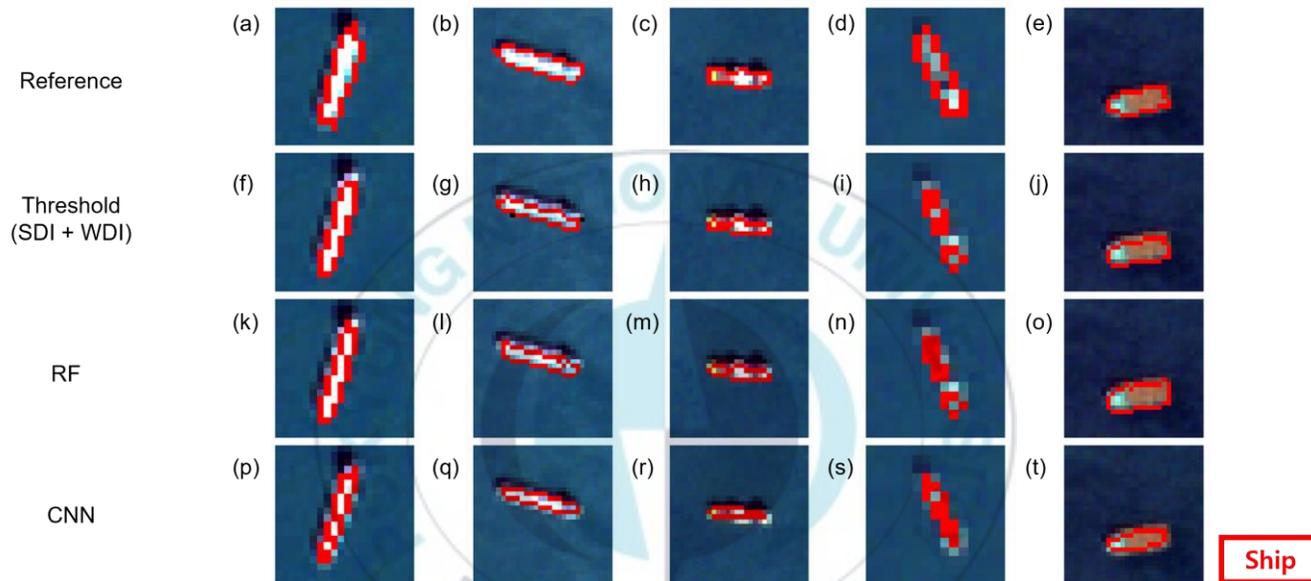


Fig. 13. Ship detection results that do not include the wake of three models.

4.2. Ship Detection Results (including wake)

Fig. 14 is a qualitative comparison of reference data and ship detection results by model for ships that include wake at the rear of the ship. Since the wake is included, the boundary between the ship and the wake is ambiguous, so there was a greater difference in the shape and area of the ship detected by model than the result of ship detection without the wake of Fig. 13. The results of threshold-based algorithm and RF ship detection showed similar results in terms of the shape and area of the ship, and compared to the reference data, the results of the two models showed a tendency for ships to be largely detected by classifying wake as ships. In the case of RF ship detection results, a phenomenon in which one ship was cut off also occurred. CNN ship detection results showed that the shape and area of the ship were the smallest among the three models, but among the three models, the ship detection results were the most similar to the reference data, and there was no phenomenon in which one ship was cut off. But, as in Fig. 13 the vessel was detected smaller than the reference data.

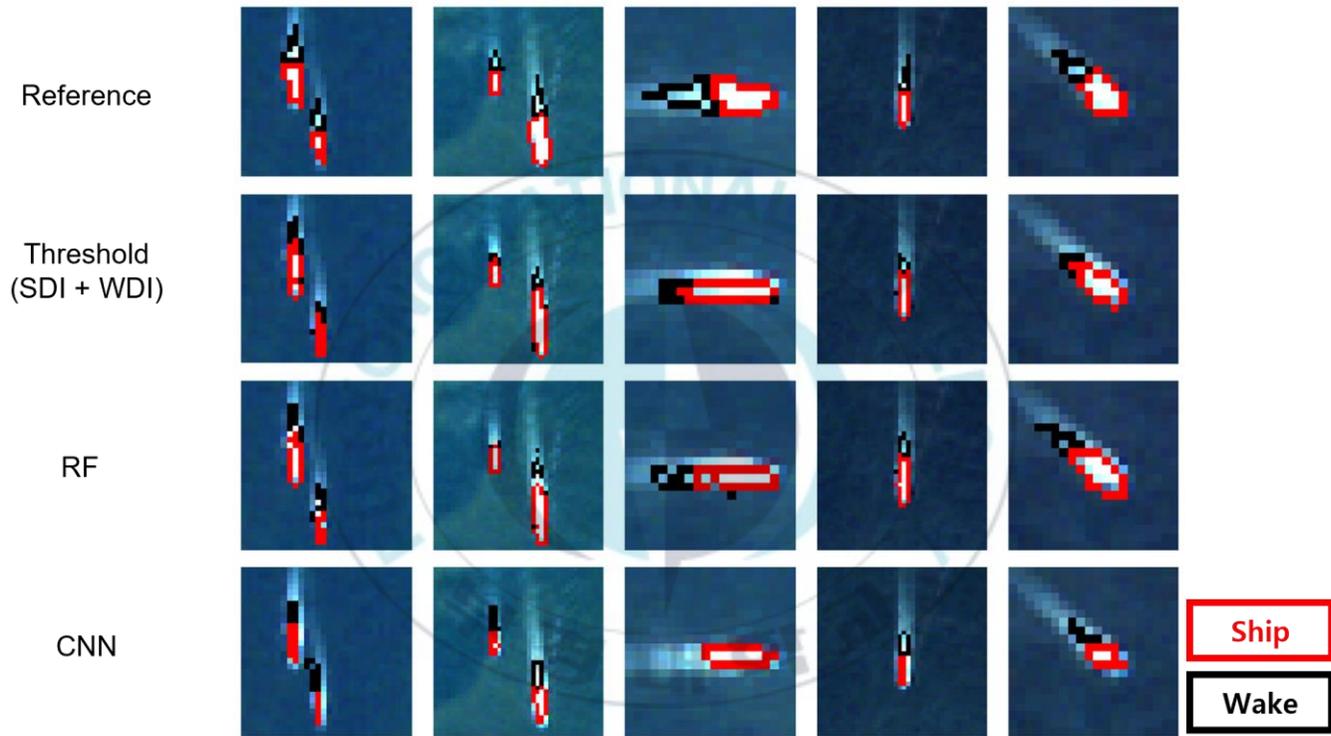


Fig. 14. Ship detection results including the wake of three models.

4.3. Combination of RF and CNN Results

To solve the phenomenon of disconnection of ships between the threshold-based algorithm and RF and the detection of ships of CNN smaller than the reference data, this study combines the RF ship detection results and the detection results of CNN ships. Fig. 15 qualitatively compares the reference data, the ship detection results of the three models, and the fusion results of RF and CNN ship detection for ships that do not contain wake. Fig. 16 qualitatively compares the reference data, the ship detection results of the three models, and the fusion results of RF and CNN ship detection for ships containing the wake. As a result of RF-CNN fusion, the area where the ship was cut off and the point where the ship was detected small have improved. Therefore, the fusion result of RF and CNN showed the shape and area of the ship most similar to the reference data.

Table. 5 shows the performance evaluation results of the three models and the combined RF and CNN results. As a result of combining RF and CNN, Recall showed the highest value of 0.97. In addition, it is significant in that the F1-score value and the IoU value-maintained accuracy similar to the other three models, even though the phenomenon of ship disconnection and the tendency of ship detection were improved.

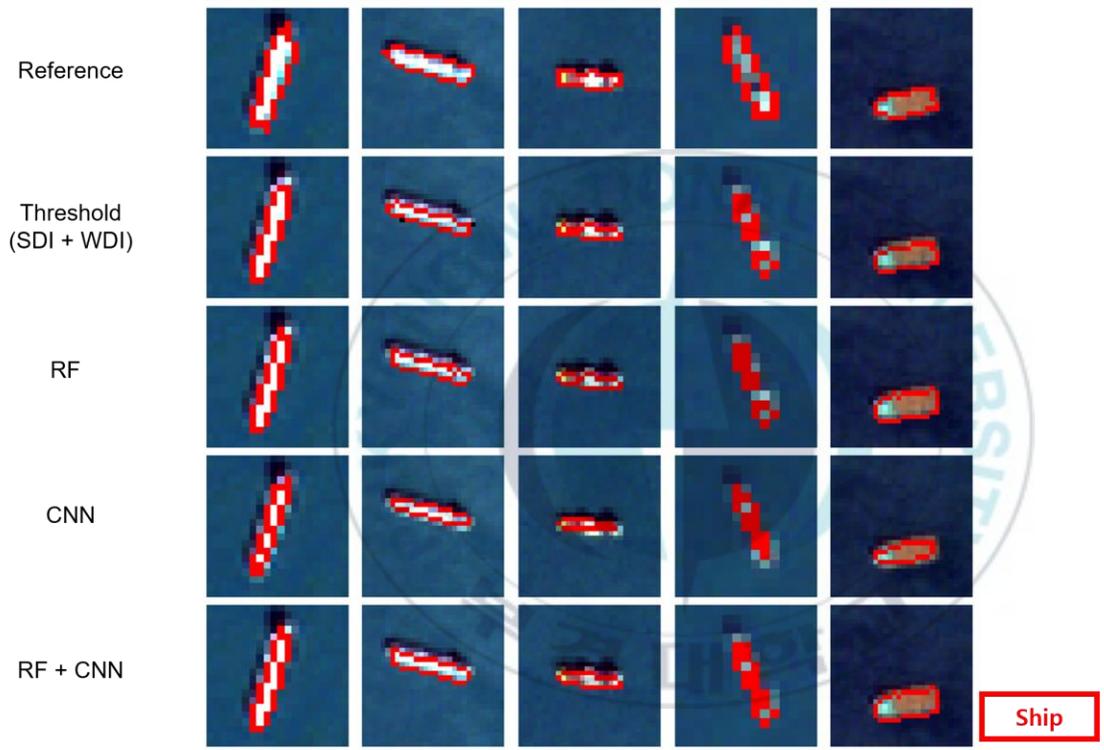


Fig. 15. Ship detection results that do not include the wake of three models and ship detection results that combine RF and CNN.

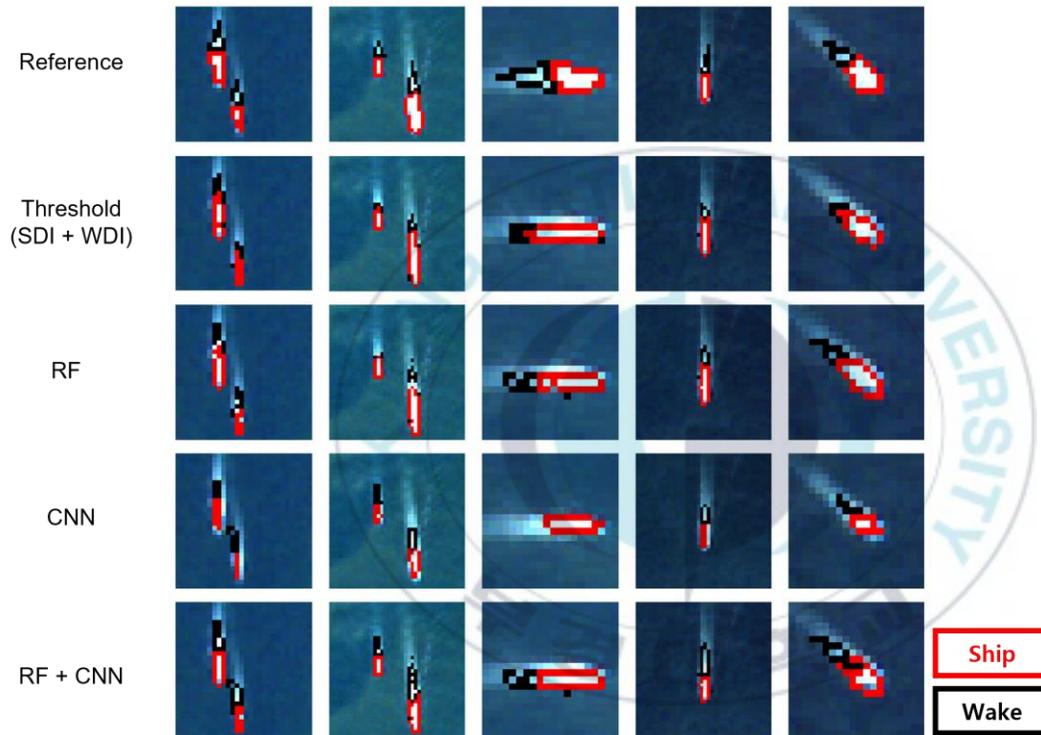
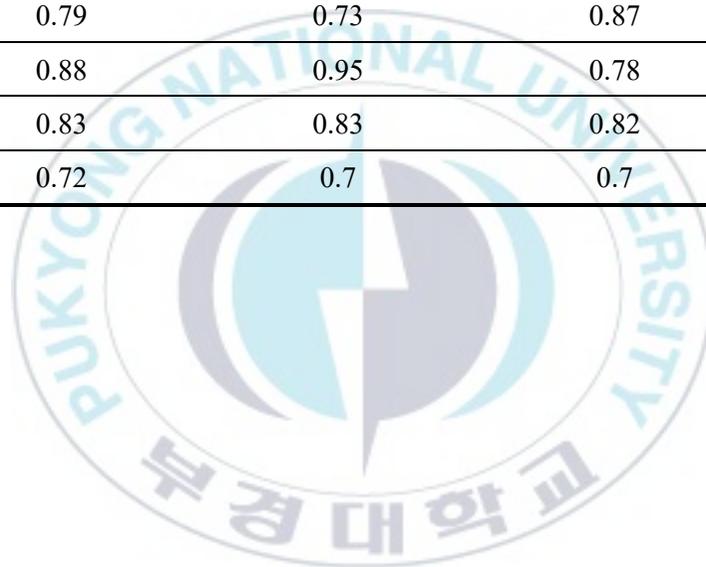


Fig. 16. Ship detection results including the wake of three models and ship detection results combining RF and CNN.

Table 5. Detection performance.

	Threshold	RF	CNN	RF + CNN
Precision	0.79	0.73	0.87	0.74
Recall	0.88	0.95	0.78	0.97
F1-score	0.83	0.83	0.82	0.84
IoU	0.72	0.7	0.7	0.72



5. Summary and conclusions

Ship detection is widely used in areas such as maritime security, maritime transportation, illegal fishing, and border control. In addition, ship detection is important to respond quickly to ship accidents that continue to occur due to an increase in maritime traffic and to minimize damage. According to a number of international regulations, ships of a particular class must be equipped with transponders that transmit the ship's ID and location at certain repetitive intervals. However, small vessels (less than 300 tons) are not obligated to carry them, and location information may not be transmitted intentionally or accidentally. Therefore, it is economical to perform ship detection using data that can be observed periodically in a wide range remotely, such as satellite images. Recently, due to the increase in optical satellites, ship detection using optical images with a large amount of data that can be used and a simple structure has been performed a lot. However, ship detection in optical images has the possibility that objects with similar brightness to ships, such as clouds and wake, may generate false alarms. Therefore, in this study, the accuracy of ship detection was improved by removing the wake, which is the main element of false alarm. Ship detection was performed by learning three models using Sentinel-2A-2B/MSI channel data. First, the Threshold-

based algorithm performed ship detection based on SDI using the ship's spectral characteristics, and the wake that were not distinguished by SDI were further analyzed to develop WDI to remove the wake that generate false alarms. Second, ships and wake detection were performed using RF with high efficiency and prediction accuracy for large-capacity dataset. Finally, among CNN techniques, an image-based object detection method, ship and wake detection were performed using Detectron2. All three models had good detection of all vessels without undetected vessels in relation to the presence or absence of the vessel. However, in the case of Threshold-based algorithms and RF, one ship was cut, distinguished for each pixel based on the spectral characteristics of each pixel. On the other hand, although it did not occur in the results of CNN, which performs object detection based on images, there was a tendency to detect ships smaller than the reference data. Therefore, in this study, the ship disconnection phenomenon and the ship small detection phenomenon were improved by fusing the RF and CNN ship detection results. The quantitative verification of the final RF-CNN fusion results showed the accuracy of Precision 0.74, Recall 0.97, F1-score 0.84, and IoU 0.72.

This study performed ship detection using three models from optical images. It is significant in that it has increased accuracy by simultaneously removing wake when performing ship detection. In addition, it is

significant in that the ship detection results were compared and analyzed by model, and the results were combined to supplement the limitations of each model while maintaining accuracy. The resolution of the optical image continues to improve. In the future, using optical satellite images with improved spatial resolution is expected to perform ship detection and wake detection with higher accuracy. It is also expected to be used for monitoring illegal China fishing boats and small North Korea's ships.



6. References

- Bouma, H., Dekker, R.J., Schoemaker, R.M., Mohamoud, A.A., 2013. Segmentation and wake removal of seafaring vessels in optical satellite images, in: *Electro-Optical Remote Sensing, Photonic Technologies, and Applications VII; and Military Applications in Hyperspectral Imaging and High Spatial Resolution Sensing*, SPIE. pp. 93–103.
- Breiman, L., 1996. Bagging predictors. *Machine learning* 24, 123–140.
- Breiman, L., 2001. Random forests. *Machine learning* 45, 5–32.
- Buck, H., Sharghi, E., Bromley, K., Guilas, C., Chheng, T., 2007. Ship detection and classification from overhead imagery, in: *Applications of Digital Image Processing XXX*, SPIE. pp. 522–536.
- Dragosevic, M.V., Vachon, P.W., 2008. Estimation of ship radial speed by adaptive processing of radarsat-1 fine mode data. *IEEE Geoscience and Remote Sensing Letters* 5, 678–682.
- Eldhuset, K., 1996. An automatic ship and ship wake detection system for spaceborne sar images in coastal regions. *IEEE transactions on Geoscience and Remote Sensing* 34, 1010–1019.
- de Figueiredo, L.P., Schmitz, T., Lunelli, R., Roisenberg, M., de Freitas, D.S., Grana, D., 2021. Direct multivariate simulation-a stepwise

- conditional transformation for multivariate geostatistical simulation. *Computers & Geosciences* 147, 104659.
- Girshick, R., 2015. Fast r-cnn, in: *Proceedings of the IEEE international conference on computer vision*, pp. 1440–1448.
- Girshick, R., Donahue, J., Darrell, T., Malik, J., 2014. Rich feature hierarchies for accurate object detection and semantic segmentation, in: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 580–587.
- He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition, in: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778.
- Heiselberg, H., 2016. A direct and fast methodology for ship recognition in sentinel-2 multispectral imagery. *Remote Sensing* 8, 1033.
- Kanjir, U., Greidanus, H., Oštir, K., 2018. Vessel detection and classification from spaceborne optical images: A literature survey. *Remote sensing of environment* 207, 1–26.
- Lan, J., Wan, L., 2009. Automatic ship target classification based on aerial images, in: *2008 International Conference on Optical Instruments and Technology: Optical Systems and Optoelectronic Instruments*, SPIE. pp. 316–325.
- Li, N., Ding, L., Zhao, H., Shi, J., Wang, D., Gong, X., 2018. Ship

- detection based on multiple features in random forest model for hyperspectral images. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci* , 891–895.
- Li, X., Chong, J., 2008. Processing of envisat alternating polarization data for vessel detection. *IEEE Geoscience and Remote Sensing Letters* 5, 271–275.
- Liu, Y., Cui, H.Y., Kuang, Z., Li, G.Q., 2017. Ship detection and classification on optical remote sensing images using deep learning, in: *ITM Web of Conferences, EDP Sciences*. p. 05012.
- Liu, Z., Zhou, F., Bai, X., Yu, X., 2013. Automatic detection of ship target and motion direction in visual images. *International Journal of Electronics* 100, 94–111.
- Mitchell, T.M., Mitchell, T.M., 1997. *Machine learning*. volume 1. McGraw-hill New York.
- Park, J.J., Oh, S., Park, K., Lee, M.S., Jang, J.C., Lee, M., 2018. A methodology of ship detection using high-resolution satellite optical image. *Journal of the Korean earth science society* 39, 241–249.
- Pham, V., Pham, C., Dang, T., 2020. Road damage detection and classification with detectron2 and faster r-cnn, in: *2020 IEEE International Conference on Big Data (Big Data)*, IEEE. pp. 5592–5601.

- Ren, S., He, K., Girshick, R., Sun, J., 2015. Faster r-cnn: Towards real-time object detection with region proposal networks. *Advances in neural information processing systems* 28.
- Russell, B.C., Torralba, A., Murphy, K.P., Freeman, W.T., 2008. Labelme: a database and web-based tool for image annotation. *International journal of computer vision* 77, 157–173.
- Simonyan, K., Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.
- Song, S.P., Cho, D.J., Park, S.G., Hong, C.E., Park, S.H., Suh, S.H., 2013. Design of transportation safety system with gps precise point positioning. *Journal of Navigation and Port Research* 37, 71–77.
- Wu, Y., Kirillov, A., Massa, F., Lo, W.Y., Girshick, R., 2019. Detectron2.
- Yang, G., Li, B., Ji, S., Gao, F., Xu, Q., 2013. Ship detection from optical satellite images based on sea surface analysis. *IEEE Geoscience and Remote Sensing Letters* 11, 641–645.
- Yang, G., Li, B., Ji, S., Gao, F., Xu, Q., 2014. Ship detection from optical satellite images based on sea surface analysis. *IEEE Geoscience and Remote Sensing Letters* 3, 641–645.
- Yang, G., Lu, Q., Gao, F., 2011. A novel ship detection method based on sea state analysis from optical imagery, in: *2011 Sixth International*

Conference on Image and Graphics, IEEE. pp. 466–471.

Zou, Z., Shi, Z., 2016. Ship detection in spaceborne optical image with svd networks. IEEE Transactions on Geoscience and Remote Sensing 54, 5832–5845.

