Geosfera: Jurnal Penelitian Geografi (GeoJPG)
Vol.5 No. 2 Desember 2026, Hal.149-161
DOI: https://doi.org/10.37905/geojpg.v5i2.40457

p - ISSN: 2962-5424
e - ISSN: 2962-5416

Oriented Object Detection and Classification of Offshore Tin Mining Vessels Using YOLOv11 in Bangka Belitung, Indonesia

Imam Purwadi[1]*, Peter D Erskine[2]

[1] Program Studi Teknik Geologi, Fakultas Teknik, Universitas Sriwijaya, Indonesia
[2] Sustainable Mineral Institute, University of Queensland, Australia
*Email Koresponden: imampurwadi@ft.unsri.ac.id

Diterima: 16-07-2026 **Disetujui: 15-10-2026** **Publish: 01-12-2026**

Abstract Offshore tin mining in Bangka Belitung, Indonesia, has created increasing challenges for coastal and marine spatial governance, particularly in distinguishing vessels associated with fishing and mining activities. Conventional monitoring approaches are often constrained by the spatial extent of offshore areas and limited access to vessel level information. This study investigates the capability of YOLOv11 based oriented object detection to identify and classify three vessel categories relevant to offshore tin mining activities, namely fishing boats, mining pontoons, and dredging vessels, using satellite basemap imagery. A dataset of 104 images containing 593 labeled vessel instances was collected from coastal areas of Bangka, Belitung, and Singkep Islands and annotated using oriented bounding boxes. Five YOLOv11 variants were fine-tuned using transfer learning and evaluated using precision, recall, mAP@50, and mAP@50–95. Among the tested models, YOLOv11l achieved the highest recall (0.971) and a mAP@50 of 0.974, while YOLOv11x obtained the highest precision (0.991) and mAP@50 (0.977). However, mAP@50–95 remained substantially lower across models, ranging from 0.616 to 0.650, indicating that precise object localization remains more challenging than detection at the IoU 0.50 threshold. The results demonstrate the feasibility of oriented object detection for distinguishing functionally different vessels in complex coastal environments. The approach provides a potential basis for satellite based spatial monitoring of offshore mining activities and can support marine spatial governance, although further validation using independent test data, temporal imagery, and field-based observations is required before operational deployment.

Keywords: Oriented Object Detection; Yolov11; Satellite Imagery; Offshore Tin Mining; Marine Spatial Monitoring

1. Introduction

Satellite based maritime surveillance has become an increasingly important approach for monitoring coastal and offshore activities, particularly where conventional vessel tracking and field-based observation are constrained by the spatial extent of marine areas and limited accessibility. Freely available and commercial remote sensing data can provide repeated observations of vessel activity and complement conventional tracking systems such as the Automatic Identification System (AIS), including the detection of vessels that do not transmit their positions, commonly referred to as "dark ships" (Heiselberg et al., 2024; Nakayama et al., 2023). Advances in satellite remote sensing have further expanded the range of maritime monitoring applications. Nighttime light imagery, for example, has been used to detect vessels under low light conditions, while high resolution satellite imagery has supported the estimation of vessel characteristics such as deadweight tonnage for maritime transportation monitoring (Zhao et al., 2023; Hou et al., 2023). The increasing availability of high-resolution imagery, combined with advances in artificial intelligence, has created opportunities to automate vessel detection and classification at spatial scales that are difficult to address through field-based surveillance alone.

Indonesia provides an important context for such applications because coastal and offshore areas are closely connected to resource extraction, fisheries, transportation, and local livelihoods. Tin mining represents one of the longstanding extractive activities in the Bangka Belitung region. Mining activities on Bangka Island have occurred since the early eighteenth century and continued through the colonial period and the post-independence development of the national tin industry, with PT Timah Tbk becoming the principal operator after mining activities were consolidated by the Indonesian government (Irzon, 2021; Ulya et al., 2024). Declining terrestrial reserves have encouraged the expansion of mining toward offshore areas, where suction dredges and cutter suction dredgers have been used as part of offshore extraction activities (Irzon, 2021). The spatial transition from terrestrial to marine mining has

Copyright © 2026 The Authors. Published by Department of Earth Science and Technology, Universitas Negeri Gorontalo
This work is licensed under a Creative Commons Attribution (CC-BY) 4.0 International License.

increased interactions among mining operations, fishing activities, coastal ecosystems, and communities whose livelihoods depend on marine resources.

Offshore mining has generated social and environmental concerns within coastal communities, particularly where mining activities overlap with fishing grounds and areas of ecological or cultural importance. Batu Beriga Village provides an example of community opposition to offshore mining because of concerns regarding coastal ecosystems, fishing livelihoods, and local cultural values (A. Putri et al., 2024; F. Putri et al., 2023). Such conflicts illustrate that offshore mining is not solely an issue of mineral extraction but also a spatial governance problem involving competing uses of marine space. Unauthorized mining activities further complicate this situation because their distribution and operational characteristics may change over time and may be difficult to monitor using conventional enforcement mechanisms. Offshore tin mining in Bangka Belitung therefore involves interconnected dimensions of spatial conflict, environmental degradation, regulatory compliance, and institutional coordination (Agustari & Muslim, 2024). Existing marine spatial management arrangements have also faced institutional challenges, including changes in authority following the implementation of Law No. 3 of 2020, which has affected the capacity of regional governments to manage and enforce marine activities (Yanto et al., 2023). These conditions increase the importance of spatially explicit monitoring approaches capable of providing information about the distribution and characteristics of vessels operating within offshore mining areas.

Remote sensing and artificial intelligence have increasingly been applied to address vessel monitoring problems. Sentinel 1 and VIIRS imagery, for example, have demonstrated the ability to detect vessel activity under conditions where conventional optical observation may be limited, including nighttime and cloudy conditions (Awak et al., 2023; Winarso et al., 2014). Deep learning approaches have subsequently expanded the potential for automated vessel detection. YOLO based models have been applied to ship detection from satellite imagery, including studies using YOLOv3 and more recent architectures (Fandisyah et al., 2021). Transfer learning using architectures such as ResNet50 and DenseNet has also demonstrated strong performance for object detection tasks involving satellite imagery (Hayati et al., 2025). More recent studies have extended YOLO based approaches to vessel classification and detection across different maritime contexts. Parasar et al. (2026), for example, applied YOLO to distinguish vessel categories such as cargo ships, tankers, naval vessels, and smaller craft from satellite imagery, while Magalhães et al. (2025) evaluated YOLOv8 and YOLOv10 for vessel and wake detection using Sentinel 2 imagery (Magalhães et al., 2025; Parasar et al., 2026).

The operational requirements of offshore tin mining surveillance, however, differ from generic ship detection tasks. Vessel monitoring for marine spatial governance requires not only locating vessels but also distinguishing vessel types that have different functional relationships with the surrounding coastal environment. Fishing boats, mining pontoons, and dredging vessels may occur within the same coastal space and can exhibit similar visual characteristics when observed from satellite imagery. Their spatial distribution nevertheless represents different forms of marine use and, consequently, carries different implications for fisheries, mining regulation, environmental management, and coastal livelihoods. Existing vessel detection studies have predominantly addressed generic ship detection or broad vessel categories, whereas function specific identification of vessels associated with offshore tin mining requires a more context sensitive classification framework. An oriented object detection approach is particularly relevant because vessels may occur at different orientations and may be located close to one another within spatially complex coastal scenes.

YOLOv11 provides an appropriate framework for examining this problem because its oriented bounding box capability allows objects to be localized according to their spatial orientation rather than being restricted to conventional horizontal bounding boxes. Such representation is relevant to satellite-based vessel detection, where the orientation, size, and spatial arrangement of individual vessels can vary substantially. A transfer learning strategy also provides a practical means of adapting pretrained detection models to a relatively specialized dataset containing vessel categories that are difficult to obtain at large scale. The present study therefore applies five YOLOv11 variants, namely YOLOv11n, YOLOv11s, YOLOv11m, YOLOv11l, and YOLOv11x, to high resolution satellite basemap imagery collected from coastal areas of Bangka, Belitung, and Singkep Islands. The analysis focuses on the detection and

Copyright © 2022 The Authors. Published by Department of Earth Science and Technology, Universitas Negeri Gorontalo
This work is licensed under a Creative Commons Attribution (CC-BY) 4.0 International License.

classification of fishing boats, mining pontoons, and dredging vessels using oriented bounding boxes, with model performance assessed through precision, recall, mAP@50, and mAP@50–95.

The resulting approach connects automated object detection with a geographically specific monitoring problem in a region where offshore mining, fisheries, environmental concerns, and marine spatial governance intersect. Rather than treating vessel detection solely as a computer vision task, the study examines its potential contribution to identifying functionally different vessel activities within a complex coastal environment. Such information can provide a spatial basis for subsequent monitoring, hotspot identification, and marine management interventions, while the limitations associated with image source, dataset size, temporal information, and independent validation remain important considerations for interpreting the applicability of the resulting models.

2. Methods

2.1 Study Area and Data Collection

The study focused on coastal areas surrounding Bangka, Belitung, and Singkep Islands, Indonesia, where offshore tin mining activities occur alongside fishing and other marine uses. These three islands were selected because they represent coastal environments where different vessel activities can occur within relatively close spatial proximity. The spatial setting provides an appropriate context for examining automated vessel detection and classification from satellite basemap imagery.

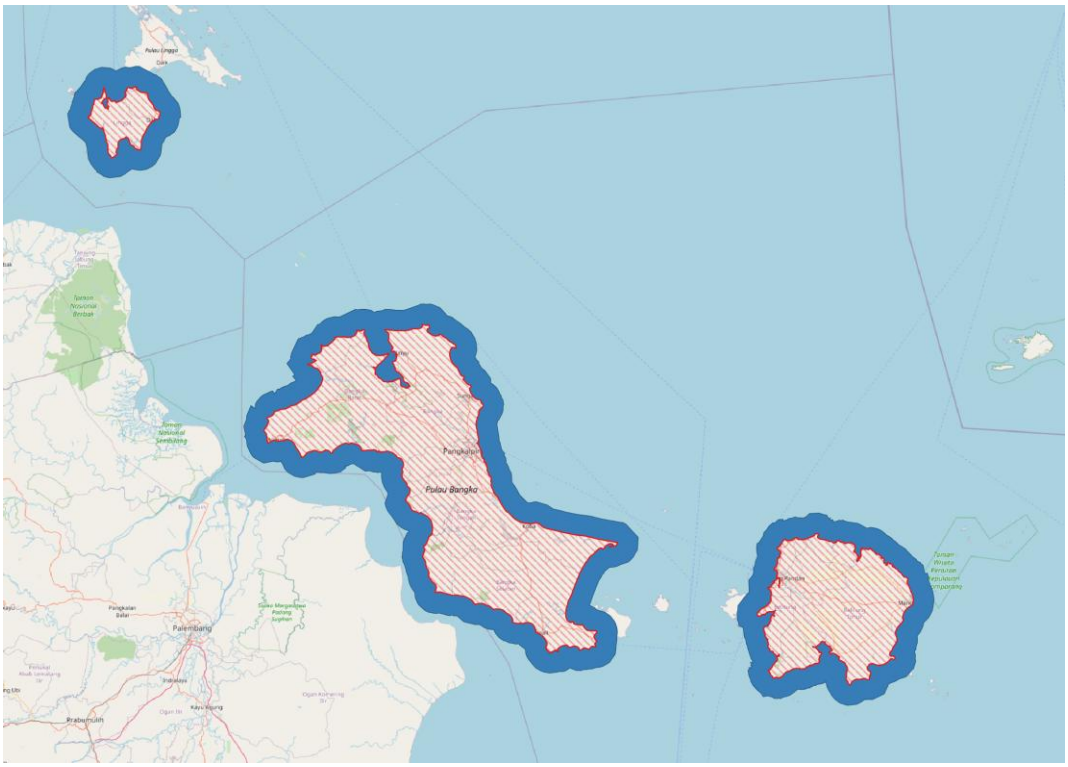


Figure 1. Coastal observation areas used in this study. The three islands marked with thick blue boundaries, from right to left, are Singkep, Bangka, and Belitung Islands. The blue boundaries represent the focus areas to locate and classify ships.

A desk-based research design was employed. Data collection, vessel identification, image selection, and annotation were conducted remotely through visual interpretation of satellite basemap imagery, without physical field surveys or vessel level ground truthing. The image collection and manual annotation were conducted in January 2025, with approximately one week allocated to the annotation process.

2.2 Data Collection and Annotation

High resolution imagery available through Google Maps was used as the primary data source. The imagery was accessed through the Google Maps interface and used to identify visible vessels within the coastal observation areas. Google Maps basemap imagery represents a composite product derived from satellite and aerial imagery from sources and acquisition dates that are not disclosed for individual scenes. Consequently, the specific sensor, acquisition date, and original imaging conditions could not be

Copyright © 2022 The Authors. Published by Department of Earth Science and Technology, Universitas Negeri Gorontalo
This work is licensed under a Creative Commons Attribution (CC-BY) 4.0 International License.

established for each image. A fixed zoom level was maintained during image collection to provide a consistent visual scale across the observations.

A manual visual survey was conducted by systematically scanning the coastal areas of Bangka, Belitung, and Singkep Islands. The map was navigated incrementally, and screenshots were captured whenever visible vessels corresponding to the target categories were identified. Three vessel categories were considered: fishing boats, mining pontoons, and dredging vessels. The original dataset identified the second category as illegal tin mining pontoons and the dredging vessels as vessels associated with authorized offshore tin mining operations.

Total of 104 image clips were collected from the three study areas. These images contained 593 vessel instances, comprising 171 fishing boats, 309 mining pontoons, and 113 dredging vessels. The relatively close occurrence of different vessel categories within some scenes allowed the dataset to represent visually complex coastal settings where multiple vessel types may be present simultaneously.

2.3 Model Architecture and Training

Each image was manually annotated using AnyLabeling, an open-source graphical annotation tool that supports oriented bounding boxes. A rotated bounding box was drawn around each visible vessel to closely represent its spatial extent and orientation. Each vessel was assigned to one of the three predefined categories according to its visual characteristics and the classification scheme used during data collection.



Figure 2. Examples of annotated ships detected in satellite imagery. Fisher boats are marked with red bounding boxes, illegal tin mining pontoons with black boxes, and dredging vessels with green boxes. In several instances, different types of ships are observed operating near one another.

The complete dataset contained 593 labeled vessel instances. The class distribution consisted of 171 fishing boats, 309 mining pontoons, and 113 dredging vessels. The annotated images were divided into training and validation subsets using an 85%/15% image level split. This produced approximately 88 training images and 16 validation images. The training subset contained 527 vessel instances, consisting of 100 dredging vessels, 141 fishing boats, and 286 mining pontoons, while the validation subset contained 66 instances, consisting of 13 dredging vessels, 30 fishing boats, and 23 mining pontoons.

No separate test set was held out in the original experimental design. Model performance was therefore assessed exclusively using the 15% validation subset. This distinction is important when

Copyright © 2022 The Authors. Published by Department of Earth Science and Technology, Universitas Negeri Gorontalo
This work is licensed under a Creative Commons Attribution (CC-BY) 4.0 International License.

interpreting the reported performance because the validation results should not be regarded as an independent external test of model generalization.

Table 1. Example of Confusion Matrix. TP (True Positive) is correct predictions, and FP (False Positive) is incorrect predictions.

Actual\Predicted	Fisher Boat	Pontoon	Dredger
Fisher Boat	TP	FP	FP
Pontoon	FP	TP	FP
Dredger	FP	FP	TP

2.4 YOLOv11 Model Architecture and Training

The study employed YOLOv11 with oriented bounding box (OBB) support to detect and classify the three vessel categories. Five model variants were evaluated: YOLOv11n, YOLOv11s, YOLOv11m, YOLOv11l, and YOLOv11x. Each variant was fine-tuned using transfer learning from pretrained weights using the annotated vessel dataset.

Model training was performed using Google Colab Pro environment powered by an NVIDIA A100-SXM4-40 GB GPU. All models were implemented using the Ultralytics framework (version 8.3.140) running on PyTorch 2.6.0. Training images were standardized to an input resolution of 1024 x 1024 pixels and trained for 50 epochs, with a fix random seed of 0 to guarantee deterministic run-to-run behavior. The automated optimizer selection (optimizer=auto) assigned the AdamW optimizer with an initial learning rate of 0.001429, momentum of 0.9, and weight decay of 0.0005. Batch sizes were dynamically allocated using the AutoBatch feature (batch=-1) at a target CUDA memory utilization of ~60% yielding batch sizes of 29 (YOLOv11n), 14 (YOLOv11s), 7 (YOLOv11m), 5 (YOLOv11l), and 3 (YOLOv11x). Data augmentation was strictly controlled to preserve the spatial characteristics of satellite basemap: random rotation was permitted up to ±15°, translation was set to 0.1, and horizontal and vertical flips were enabled at a 0.5 probability. Scaling, shearing, perspective transformations, mosaic, mixup, and color adjustments (HSV) were disabled.

The restricted augmentation configuration was adopted to limit the introduction of synthetic variation and to retain the spatial characteristics present in the original basemap imagery. This configuration was particularly relevant because vessel appearance in the dataset was influenced by viewing orientation, vessel size, surrounding water conditions, and the spatial context of coastal scenes.

2.5 Model Evaluation

Model performance was evaluated using four standard object detection metrics: precision, recall, mAP@50, and mAP@50–95. Precision measures the proportion of predicted vessel detections that correspond to correct detections, whereas recall represents the proportion of ground truth vessels successfully detected by the model.

The mAP@50 metric represents mean average precision at an Intersection over Union (IoU) threshold of 0.50. The mAP@50–95 metric represents the mean average precision calculated across IoU thresholds from 0.50 to 0.95. The latter provides a more stringent assessment because successful detection requires increasingly precise overlap between predicted and ground truth bounding boxes.

Evaluation was conducted using the 16-image validation subset containing 66 labeled vessel instances. Model performance was compared across the five YOLOv11 variants to identify differences in detection accuracy and the trade-off between model size and predictive performance.

2.6 Classification and Error Analysis

The detection results were examined according to the three vessel categories to assess the ability of the models to distinguish visually different vessel types. Confusion matrix analysis was used to distinguish correct predictions from classification errors among fishing boats, mining pontoons, and dredging vessels. The original manuscript provides the conceptual structure of a three-class confusion matrix, in which true positives represent correctly predicted categories and false positives represent incorrect predictions.

Per class analysis was used to identify differences in model behavior among vessel categories. Attention was given to the ability of the models to detect mining related vessels and to the differences between precision and recall across the YOLOv11 variants. The analysis also considered the difference

Copyright © 2022 The Authors. Published by Department of Earth Science and Technology, Universitas Negeri Gorontalo
This work is licensed under a Creative Commons Attribution (CC-BY) 4.0 International License.

between mAP@50 and mAP@50–95 as an indicator of detection performance under less stringent and more stringent localization requirements.

3. Results and Discussions

3.1 Dataset Characteristics and Vessel Distribution

A total of 104 image clips were collected from coastal areas surrounding Bangka, Belitung, and Singkep Islands through systematic visual interpretation of high-resolution satellite basemap imagery. The image collection produced 593 annotated vessel instances representing three target categories, namely fishing boats, mining pontoons, and dredging vessels. Mining pontoons constituted the largest class, with 309 instances (52.1%), followed by fishing boats with 171 instances (28.8%) and dredging vessels with 113 instances (19.1%). The resulting distribution reflects the vessels that could be identified and annotated within the selected observation scenes rather than the actual composition of the offshore vessel population across the study region. This distinction is important because the dataset was constructed specifically for model development and evaluation and was not designed as a statistically representative survey of vessel abundance.

The composition of the dataset provides several characteristics that are relevant to the detection task. Mining pontoons account for more than half of all annotated instances, whereas dredging vessels form the smallest class. Such an unequal distribution provides the model with substantially more examples of mining pontoons during training than of dredging vessels. Class frequency may therefore influence the extent to which each category is represented in the learned visual features and subsequently affect class specific detection performance. Interpretation of model performance consequently needs to consider not only the aggregate detection metrics but also the distribution of observations available for each vessel category.

Table 2. Distribution of Vessel Instances Across the Training and Validation Datasets

Vessel category	Training	Validation	Total	Proportion (%)
Mining pontoons	286	23	309	52.1
Fishing boats	141	30	171	28.8
Dredging vessels	100	13	113	19.1
Total	527	66	593	100.0

The image dataset was divided at the image level into an approximately 85% training subset and 15% validation subset. The training data consisted of approximately 88 images containing 527 vessel instances, whereas 16 images containing 66 vessel instances were retained for validation. Within the training subset, mining pontoons represented 286 instances, fishing boats 141 instances, and dredging vessels 100 instances. The validation subset contained 23 mining pontoons, 30 fishing boats, and 13 dredging vessels. The image level division is appropriate for scenes containing multiple vessels because it keeps all objects belonging to a particular image within the same subset, avoiding the direct distribution of visually identical scene information between training and validation data.

Several characteristics of the validation data require careful consideration when interpreting the subsequent model results. Only 16 images and 66 vessel instances were available for validation, meaning that a relatively small number of prediction errors can produce a noticeable change in precision, recall, and mean average precision. The validation subset also differs in class composition from the training data. Fishing boats represent the largest class within the validation subset, whereas mining pontoons dominate the training data. Such differences are relevant when comparing class specific performance because aggregate metrics may conceal variations in model behavior among vessel categories.

No separate independent test dataset was retained in the experimental design. Model performance reported in the following sections therefore represents performance on the held-out validation subset rather than performance on an independent test dataset. This distinction limits the extent to which the reported metrics can be interpreted as evidence of generalization to previously unseen geographic areas, different image acquisition conditions, or vessel populations. The high-performance values should consequently be understood within the specific spatial and visual conditions represented by the collected dataset rather than as an unconditional estimate of operational accuracy.

Copyright © 2022 The Authors. Published by Department of Earth Science and Technology, Universitas Negeri Gorontalo
This work is licensed under a Creative Commons Attribution (CC-BY) 4.0 International License.

Image characteristics introduce another consideration. Google Maps basemap imagery was used as the primary source, and a fixed zoom level was maintained during image collection to provide a relatively consistent visual scale. Individual basemap scenes may nevertheless originate from different satellite or aerial imaging sources, while their exact acquisition dates, sensors, and imaging conditions are not consistently disclosed through the interface. The dataset may therefore contain variations in spatial detail and visual appearance that cannot be fully controlled through the image collection procedure. These characteristics are particularly relevant for small vessel detection, where object dimensions, image resolution, orientation, and surrounding environmental features can affect the distinction between vessel categories.

The spatial arrangement of vessels within individual scenes also contributes to the difficulty of the detection task. Fishing boats, mining pontoons, and dredging vessels may occur within the same coastal environment and, in some scenes, appear relatively close to one another. Such spatial configurations require the model to identify individual objects while simultaneously assigning the appropriate class and estimating their orientation. The use of oriented bounding boxes provides a representation that is better suited to elongated and variably oriented vessels than conventional horizontal bounding boxes. This characteristic becomes particularly important when several vessels are located close together because more precise representation of object orientation can help distinguish adjacent instances.

The dataset consequently provides a focused experimental basis for examining YOLOv11 oriented object detection within a coastal environment characterized by multiple vessel activities. Its principal strength is the inclusion of three vessel categories that are spatially relevant to offshore tin mining and may coexist within the same marine space. At the same time, the unequal class distribution, limited number of images, small validation subset, uncertain provenance of individual basemap scenes, and absence of an independent test set constrain the interpretation of the resulting performance metrics. These characteristics are particularly important for evaluating whether differences among YOLOv11 variants represent consistent improvements in vessel detection or are partly influenced by the composition and scale of the available dataset

3.2 YOLOv11 Model Performance

Five YOLOv11 variants with oriented bounding box support were fine-tuned using the same training dataset and experimental configuration to examine how model scale influenced vessel detection and classification performance. The evaluated variants consisted of YOLOv11n, YOLOv11s, YOLOv11m, YOLOv11l, and YOLOv11x. All models were trained for 50 epochs and evaluated using the same 16 image validation subset containing 66 labeled vessel instances. Performance was assessed using precision, recall, mAP@50, and mAP@50–95, allowing the comparison to consider both the reliability of positive detections and the ability to identify and localize the annotated vessels.

Table 2. Performance Metrics on Validation Dataset

Model	Box Precision	Recall	mAP@50	mAP@50–95
YOLOv11n	0.895	0.905	0.964	0.624
YOLOv11s	0.954	0.887	0.948	0.650
YOLOv11m	0.894	0.886	0.946	0.616
YOLOv11l	0.919	0.971	0.974	0.634
YOLOv11x	0.991	0.866	0.977	0.639

Clear differences emerged among the five variants despite their evaluation under the same dataset and training conditions. YOLOv11x produced the highest precision (0.991) and mAP@50 (0.977), indicating that its predictions were highly reliable at the IoU 0.50 threshold. YOLOv11l, however, achieved the highest recall (0.971) and a mAP@50 of 0.974, only slightly below the value obtained by YOLOv11x. The smallest YOLOv11n model also performed competitively, obtaining a precision of 0.895, recall of 0.905, and mAP@50 of 0.964. YOLOv11s recorded the highest mAP@50–95 among the five variants at 0.650, despite having a lower mAP@50 of 0.948. YOLOv11m produced the lowest mAP@50 (0.946) and mAP@50–95 (0.616), while its precision and recall were also lower than those of the smaller YOLOv11s and YOLOv11n models.

The results indicate that increasing model size did not produce a monotonic improvement across all evaluation metrics. YOLOv11x, the largest evaluated variant, achieved the highest precision and mAP@50 but did not achieve the highest recall. Its recall of 0.866 was lower than that of YOLOv11n

Copyright © 2022 The Authors. Published by Department of Earth Science and Technology, Universitas Negeri Gorontalo
This work is licensed under a Creative Commons Attribution (CC-BY) 4.0 International License.

(0.905) and substantially lower than YOLOv11l (0.971). This distinction is important for the present application because vessel surveillance requires not only accurate positive predictions but also the ability to identify as many relevant vessels as possible. A model that produces very precise detections but fails to identify a greater proportion of actual vessels may be less desirable when the monitoring objective emphasizes comprehensive spatial identification.

YOLOv11l provides a more balanced performance profile under the present validation setting. Its recall of 0.971 was the highest among the evaluated models, while its precision of 0.919 remained relatively high. The model also achieved a mAP@50 of 0.974, only 0.003 below YOLOv11x. The difference in mAP@50 between these two models is therefore very small, whereas the difference in recall is considerably larger, with YOLOv11l exceeding YOLOv11x by 0.105. This pattern suggests that the additional model capacity of YOLOv11x did not translate into greater sensitivity to all vessel instances within the validation scenes. Rather, its advantage was more apparent in reducing incorrect detections and maintaining high precision at the 0.50 IoU threshold.

An important feature of the results is the substantially lower mAP@50–95 values compared with mAP@50 across all models. The values ranged from 0.616 for YOLOv11m to 0.650 for YOLOv11s, whereas mAP@50 ranged from 0.946 to 0.977. This difference indicates that the models were highly successful at identifying vessels when a relatively permissive IoU threshold of 0.50 was applied, but their performance decreased when progressively stricter localization criteria were imposed. Such a pattern is relevant to oriented vessel detection because small vessels occupy relatively limited areas within the imagery, making small deviations in predicted object boundaries capable of producing substantial changes in IoU. High mAP@50 should therefore not be interpreted as evidence that the predicted oriented bounding boxes consistently match the ground truth boundaries with very high spatial precision.

The relationship between mAP@50 and mAP@50–95 also changes the interpretation of model ranking. YOLOv11x was strongest according to mAP@50, whereas YOLOv11s produced the highest mAP@50–95. YOLOv11l occupied an intermediate position with 0.974 mAP@50 and 0.634 mAP@50–95. Consequently, selecting a model solely according to one metric would produce different conclusions. For a monitoring application in which missing vessels is considered particularly undesirable, recall provides an important criterion and favors YOLOv11l. For applications prioritizing the minimization of false positive detections, precision favors YOLOv11x. The selection of YOLOv11l is therefore more appropriately understood as a balance among detection sensitivity, precision, and computational requirements rather than as the model that achieves the highest value for every metric.

Table 3. Training Time and Cost per Model

Model	Training Time (hrs)	Compute Units Used	Estimated Cost (USD)
YOLOv11n	0.035	0.24	$0.02
YOLOv11s	0.029	0.20	$0.02
YOLOv11m	0.047	0.32	$0.03
YOLOv11l	0.070	0.48	$0.05
YOLOv11x	0.117	0.80	$0.08

Computational requirements increased with model scale, although the relationship was not strictly reflected in detection performance. YOLOv11x required approximately 0.117 hours of training and 0.80 Compute Units, compared with 0.070 hours and 0.48 Compute Units for YOLOv11l. The estimated training costs were US$0.08 and US$0.05, respectively. All five model variants together required less than US$0.20 under the reported Google Colab Pro configuration.

The relatively small difference in detection performance between YOLOv11l and YOLOv11x becomes more meaningful when considered alongside their computational requirements. YOLOv11x improved precision from 0.919 to 0.991 and mAP@50 from 0.974 to 0.977, but its recall decreased from 0.971 to 0.866. At the same time, its training time increased from 0.070 to 0.117 hours, and its estimated cost increased from US$0.05 to US$0.08. YOLOv11l therefore provides a more favorable balance for the specific objective of identifying vessel instances in the study area, particularly because its recall advantage is considerably larger than the marginal mAP@50 advantage obtained by YOLOv11x.

The comparison should nevertheless be interpreted within the experimental conditions of this study. All models were trained and evaluated using the same relatively small dataset and the same validation

Copyright © 2022 The Authors. Published by Department of Earth Science and Technology, Universitas Negeri Gorontalo
This work is licensed under a Creative Commons Attribution (CC-BY) 4.0 International License.

subset, and no independent test dataset was available. Differences of a few percentage points between variants may therefore not represent stable differences in performance under other datasets or geographic conditions. The results are more appropriately used to identify the strongest candidate among the tested configurations within the present experimental setting rather than to establish a universal ranking of YOLOv11 architectures.

3.3 Detection Performance and Classification Patterns

Quantitative evaluation provides evidence of the overall performance of the five YOLOv11 variants, while visual inspection of the predicted oriented bounding boxes provides an additional perspective on how the models respond to vessels with different sizes, orientations, and spatial arrangements. Figure 3 presents representative validation images in which ground truth annotations are compared with the oriented bounding boxes generated by the trained models. The examples include small fishing boats as well as larger dredging vessels and scenes containing multiple vessel categories within relatively close proximity.

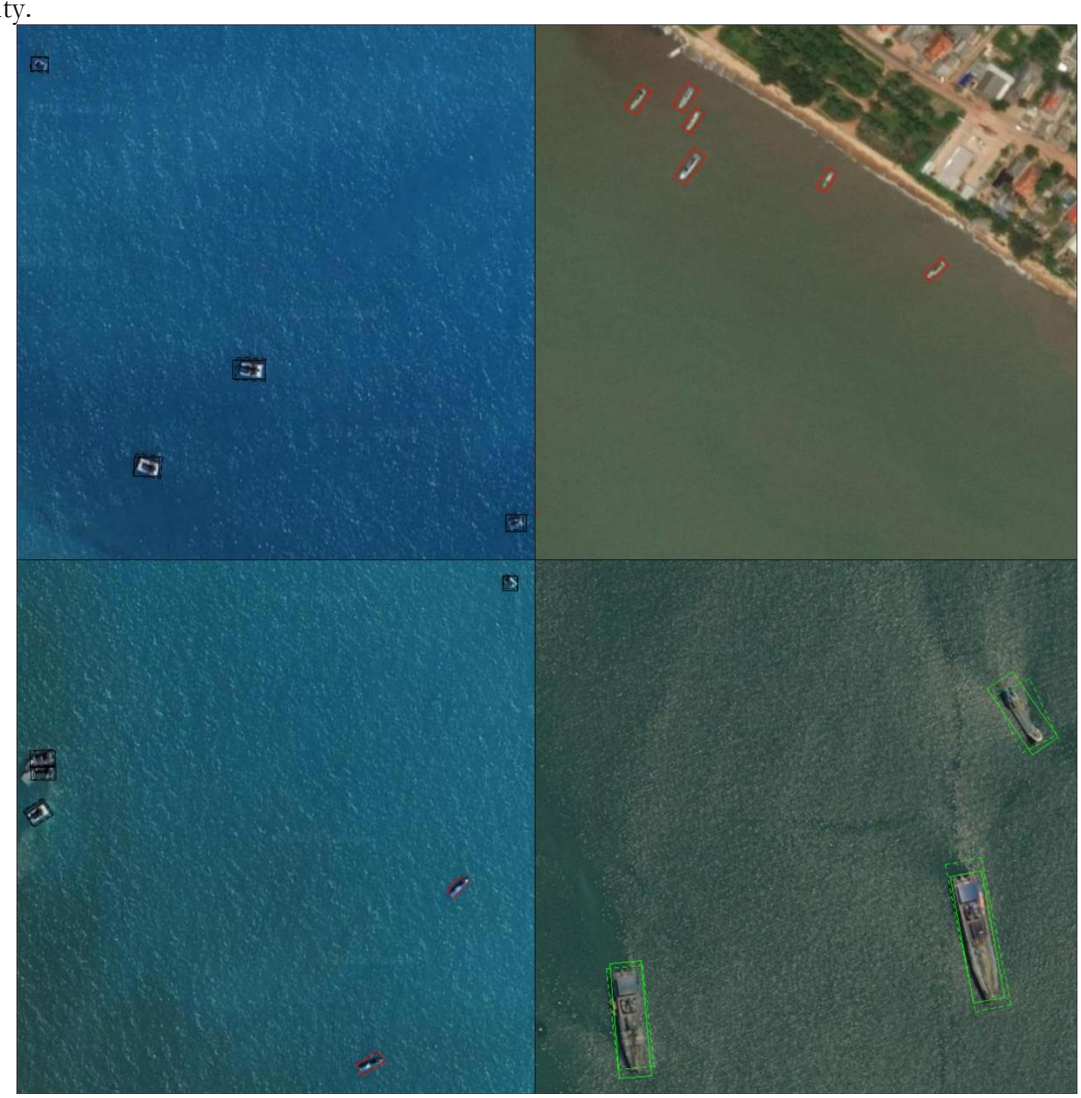


Figure 3. Examples of oriented bounding box predictions for vessels in the validation imagery. Solid boxes represent ground truth annotations, whereas dashed boxes represent model predictions. Different colors indicate the vessel categories used in the study.

Visual correspondence between the ground truth and predicted boxes indicates that the models can identify vessels across different orientations and object scales within the validation scenes. The use of oriented bounding boxes is particularly relevant because vessels are elongated objects whose longitudinal axes may vary substantially relative to the image coordinate system. A horizontal bounding box may include a considerable amount of surrounding water when a vessel is strongly rotated, whereas an oriented representation can follow the approximate geometry of the vessel more closely. The examples

Copyright © 2022 The Authors. Published by Department of Earth Science and Technology, Universitas Negeri Gorontalo
This work is licensed under a Creative Commons Attribution (CC-BY) 4.0 International License.

in Figure 3 therefore provide qualitative support for the quantitative results reported in Table 2, particularly the high mAP@50 values obtained by the evaluated models.

Several visual characteristics nevertheless make vessel classification more demanding than simple object localization. Fishing boats, mining pontoons, and dredging vessels differ in size and structural configuration, but their appearance can overlap when viewed from satellite imagery at relatively small scales. Small fishing boats may contain limited pixels from which discriminative features can be learned, while mining pontoons and dredging vessels may contain structural elements that are difficult to distinguish depending on viewing orientation and image quality. Vessels positioned close to one another introduce an additional challenge because the model must separate adjacent objects and assign each detection to the appropriate category. These conditions provide a plausible explanation for why strong detection performance at mAP@50 does not necessarily translate into equally strong performance under the stricter mAP@50–95 criterion.

The quantitative results further indicate that model performance is not uniform across evaluation dimensions. YOLOv11x achieved the highest precision (0.991) and mAP@50 (0.977), suggesting that its predictions were highly reliable when evaluated at an IoU threshold of 0.50. YOLOv11l, however, achieved the highest recall (0.971), indicating that it identified a larger proportion of the annotated vessel instances. YOLOv11s produced the highest mAP@50–95 (0.650), despite recording a lower mAP@50 of 0.948. These differences demonstrate that the ranking of the models depends on the criterion used to define detection quality rather than on model size alone.

The contrast between mAP@50 and mAP@50–95 is particularly informative for interpreting the oriented detection results. All five models produced mAP@50 values above 0.94, whereas mAP@50–95 values remained between 0.616 and 0.650. A high mAP@50 indicates that the predicted boxes generally overlap sufficiently with the annotated objects at the 0.50 IoU threshold. The substantially lower mAP@50–95 indicates that maintaining accurate overlap as the IoU requirement becomes progressively stricter is more difficult. For small and elongated vessels, even relatively minor deviations in the predicted position, dimensions, or rotation of the bounding box can reduce IoU substantially. The difference between the two metrics therefore suggests that the principal challenge is not simply recognizing that a vessel is present but achieving highly precise spatial delineation of its boundaries.

Visual interpretation also needs to be distinguished from categorical verification. The current dataset was created through manual interpretation of satellite basemap imagery, and the annotation process assigned vessels to three predefined categories. No vessel level field observations or independent operational records were available to verify every visual classification. Consequently, the model results demonstrate consistency with the annotation scheme used to construct the dataset, but they should not be interpreted as independent confirmation of the actual operational status of each vessel. This distinction is particularly important for the category of mining pontoons, where the designation of legal or illegal activity cannot be established solely from the visual appearance of a vessel in satellite imagery.

The absence of an experimental confusion matrix with class specific numerical results also limits the extent to which classification errors can be quantified in the present analysis. The original manuscript includes a three-class confusion matrix as an illustrative example rather than as an empirical result. For a stronger final version, an actual confusion matrix generated from the validation predictions should replace that illustrative table. Such a matrix would allow the analysis to identify which vessel categories are most frequently confused and whether errors are concentrated in classes. Without those empirical counts, claims regarding a specific pair of classes being the dominant source of misclassification would not be adequately supported.

The visual and quantitative findings indicate that oriented object detection can identify the vessel categories represented in the validation dataset, including scenes with different vessel orientations and multiple objects. The results also show that detection success should not be equated with precise localization or independent verification of vessel activity. This distinction is important for a geographically oriented application because the ultimate value of vessel detection lies not only in producing a high detection score but also in generating spatial information that can subsequently be combined with other evidence for monitoring offshore activities.

Copyright © 2022 The Authors. Published by Department of Earth Science and Technology, Universitas Negeri Gorontalo
This work is licensed under a Creative Commons Attribution (CC-BY) 4.0 International License.

3.4 Implications and Limitations for Offshore Tin Mining Monitoring

The results demonstrate that satellite image based oriented object detection can provide a practical means of distinguishing vessel categories that have different functional relationships with offshore activities in the Bangka Belitung and Singkep coastal region. Unlike generic ship detection, which primarily addresses whether a vessel is present within an image, the present classification framework differentiates fishing boats, mining pontoons, and dredging vessels. This distinction is particularly relevant in coastal areas where fishing and mining activities occur within interconnected marine spaces. Identification of vessel categories can provide more meaningful spatial information for subsequent assessment of marine use than undifferentiated vessel counts because the presence and concentration of different vessel types may represent different patterns of resource utilization.

The geographical relevance of this approach lies in its potential to transform vessel detection from an isolated image interpretation task into a source of spatial information for marine monitoring. Vessel locations derived from satellite imagery can potentially be integrated with other spatial datasets to examine the distribution of mining related activities, fishing areas, coastal settlements, environmental sensitive zones, and marine spatial planning boundaries. Such integration could support the identification of areas where different marine uses overlap and provide an additional evidence layer for monitoring coastal activities. Previous research has emphasized the environmental and governance challenges associated with offshore tin mining in Bangka Belitung, including conflicts arising from competing marine uses and limitations in the implementation of marine spatial regulations (Agustari & Muslim, 2024). The vessel classification approach developed here therefore has relevance beyond object recognition because it provides an initial mechanism for generating spatially explicit information about the types of activities represented by visible vessels.

Potential application to offshore mining monitoring should nevertheless be understood as a supporting function rather than a replacement for existing regulatory and field-based systems. A satellite image can indicate the presence, location, orientation, and apparent visual characteristics of a vessel, but these observations alone do not establish whether a particular mining activity is legally authorized. Legal status depends on information that is not necessarily visible in an image, including permits, operational authorization, concession boundaries, vessel registration, and temporal validity. The same limitation applies to the distinction between authorized and unauthorized mining activities. Consequently, the identification of a mining pontoon from its visual characteristics should not be interpreted as direct evidence of illegal activity without corroborating administrative, spatial, or field based information.

Integration with vessel tracking and spatial regulatory data would substantially strengthen the practical value of the approach. AIS and VMS can provide complementary information on vessel identity, movement, and temporal activity, while satellite imagery can provide independent visual evidence of vessels that may not be adequately represented in tracking systems. Combining these data sources could allow future studies to examine whether detected vessels correspond to registered operations, whether their observed positions fall within designated mining or fishing areas, and whether their spatial distribution changes over time. The original study identifies temporal tracking and integration with AIS or VMS as important directions for extending the approach toward dynamic monitoring.

Several methodological limitations constrain the interpretation and broader application of the findings. The first concerns dataset size and composition. Only 104 image clips containing 593 vessel instances were available, and the three classes were unevenly represented, with mining pontoons accounting for 309 instances, fishing boats for 171, and dredging vessels for 113. The validation subset contained only 16 images and 66 vessel instances. Such a limited evaluation sample makes the reported metrics sensitive to individual prediction errors and does not provide sufficient evidence to establish model performance across the full range of coastal conditions encountered in Indonesia.

The second limitation concerns generalization. No independent test dataset was retained in the experimental design, so the reported performance represents validation performance rather than an independent assessment of generalization. The models were evaluated using images drawn from the same geographically defined observation framework used to construct the dataset. Their ability to maintain comparable performance under different geographic settings, image acquisition conditions, vessel densities, or seasonal conditions therefore remains untested. This limitation is particularly important for

Copyright © 2022 The Authors. Published by Department of Earth Science and Technology, Universitas Negeri Gorontalo
This work is licensed under a Creative Commons Attribution (CC-BY) 4.0 International License.

remote sensing-based detection because changes in spatial resolution, image quality, vessel orientation, environmental background, and object size may affect the visual features learned by the model.

Image provenance presents a further limitation. Google Maps basemap imagery was selected because it provides accessible high resolution visual data, but the underlying basemap is a composite product for which the exact sensor and acquisition date of individual images cannot consistently be determined. This characteristic limits temporal reproducibility and makes it difficult to establish whether observations from different locations represent the same temporal condition. The resulting model should therefore be regarded as a proof of feasibility under the available imagery conditions rather than a fully validated multi-temporal monitoring system.

Absence of field-based vessel level ground truth also requires particular attention. Vessel identification and annotation were conducted through remote visual interpretation, without physical verification of individual vessels. The evaluation consequently measures the model's agreement with the manually constructed annotation dataset. It does not independently establish the actual identity, operational function, or legal status of every vessel observed in the imagery. Future validation should incorporate field observations, vessel registries, AIS/VMS records, or other authoritative sources to strengthen the reliability of the classification labels and provide an external basis for evaluating model predictions.

These limitations define the appropriate scope of the study's contribution. The findings support the feasibility of using YOLOv11 with oriented bounding boxes to generate vessel level spatial information from accessible satellite basemap imagery, particularly for differentiating vessel categories relevant to offshore tin mining. They do not yet establish a fully operational enforcement system, a real time monitoring platform, or a generalized classifier for Indonesian maritime environments. Future research should therefore prioritize larger and more balanced datasets, independent geographic and temporal test sets, field-based validation, and integration with AIS or VMS. Extending the framework toward multi-temporal analysis would also enable vessel distributions to be examined as dynamic spatial patterns rather than as isolated observations, providing a stronger basis for marine spatial planning, environmental monitoring, and evidence-based management of offshore tin mining activities.

4. Conclusions

This study demonstrates the feasibility of using YOLOv11 with oriented bounding boxes to detect and classify fishing boats, mining pontoons, and dredging vessels from satellite imagery in coastal areas surrounding Bangka, Belitung, and Singkep Islands. Among the evaluated models, YOLOv11l provided the most balanced performance, achieving a recall of 0.971 and mAP@50 of 0.974, while requiring lower computational resources than YOLOv11x. The results indicate that vessel detection can provide useful spatial information for supporting offshore tin mining monitoring and marine spatial management when integrated with complementary data such as AIS, VMS, and field observations.

Future implementation should prioritize the development of larger and independently validated datasets, including multi-temporal imagery and field-based verification, to improve the reliability and spatial transferability of vessel detection. Integration of the resulting vessel information with AIS/VMS data and marine spatial planning datasets is also recommended to strengthen the monitoring of vessel activities and support evidence-based management of offshore tin mining areas.

REFERENSI

Agustari, A., & Muslim, A. (2024). Minimizing Environmental Conflict Due to Offshore Tin Mining (Study of the Implementation of Regional Regulation RZWP3K in Bangka Belitung). *Jurnal Ekonomi Dan Kebijakan Publik*, *14*(2), 171–190. https://doi.org/10.22212/jekp.v14i2.3073

Awak, D. S. H. L. M. K., Lumban-Gaol, J., & Kushardono, D. (2023). Deteksi Kapal Penangkapan Ikan Menggunakan data Visible Infrared Imaging Radiometer Suite (VIIRS) dan data Vessel Monitoring System (VMS) di Wilayah Pengelolaan Perikanan Negara Republik Indonesia. *Journal of Marine and Aquatic Sciences*, *8*(1), 102. https://doi.org/10.24843/jmas.2022.v08.i01.p12

Fandisyah, A. F., Iriawan, N., & Winahju, W. S. (2021). Deteksi Kapal di Laut Indonesia Menggunakan YOLOv3. *Jurnal Sains Dan Seni ITS*, *10*(1). https://doi.org/10.12962/j23373520.v10i1.59312

Hayati, N., Qolbi, S. N., Utama, I. K. A. P., Putranto, T., Satrio, D., & Cahyadi, M. N. (2025). Ship detection based on variety of YOLO using multi temporal and polarization SAR images. *Earth Science Informatics*, *18*(1), 155. https://doi.org/10.1007/s12145-024-01503-3

Copyright © 2022 The Authors. Published by Department of Earth Science and Technology, Universitas Negeri Gorontalo
This work is licensed under a Creative Commons Attribution (CC-BY) 4.0 International License.

Heiselberg, P., Pedersen, H. B., Sørensen, K. A., & Heiselberg, H. (2024). Identification of Ships in Satellite Images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *17*, 6045–6054. https://doi.org/10.1109/JSTARS.2024.3368508

Hou, M., Li, Y., Xie, M., Wang, S., & Wang, T. (2023). Monitoring vessel deadweight tonnage for maritime transportation surveillance using high resolution satellite image. *Ocean & Coastal Management*, *239*, 106607. https://doi.org/10.1016/j.ocecoaman.2023.106607

Irzon, R. (2021). Penambangan timah di Indonesia: Sejarah, masa kini, dan prospeksi. *Jurnal Teknologi Mineral Dan Batubara*, *17*(3), 179–189. https://doi.org/10.30556/jtmb.Vol17.No3.2021.1183

Magalhães, R., Falcão, A. P., & Barbosa, A. (2025). Vessel detection leveraging satellite imagery and YOLO in maritime surveillance. *Remote Sensing Applications: Society and Environment*, *40*, 101730. https://doi.org/10.1016/j.rsase.2025.101730

Nakayama, S., Dong, W., Correro, R. G., Selig, E. R., Wabnitz, C. C. C., Hastie, T. J., Leape, J., Yeung, S., & Micheli, F. (2023). Comparing spatial patterns of marine vessels between vessel-tracking data and satellite imagery. *Frontiers in Marine Science*, *9*. https://doi.org/10.3389/fmars.2022.1076775

Parasar, D., Kukreja, S., Dashpute, T., Dubey, S., Bhinde, S., & Singh, H. (2026). Ship Classification Using Satellite Imagery: A YOLO-Based Approach. In *Lecture Notes in Networks and Systems: 1982 LNNS* (pp. 391–401). Springer Science and Business Media Deutschland GmbH. https://doi.org/10.1007/978-3-032-26376-6_35

Putri, A., Putri Cahyani, Y., Fariza Yulianti Saputri, A., Hukum, F., & Hukum, J. (2024). Resistensi Masyarakat Terhadap Pertambangan Timah Di Laut (Studi Kasus: Desa Batu Beriga, Provinsi Kepulauan Bangka Belitung). *BULLET : Jurnal Multidisiplin Ilmu*, *3*, 227–232. https://journal.mediapublikasi.id/index.php/bullet

Putri, F., Valensia, M., Purnama, R., & Manik, J. (2023). Dampak Kerusakan Lingkungan Biotik, Abiotik, dan Sosial Budaya Akibat Pertambangan Timah Ilegal di Kecamatan Mentok. *SENTRI: Jurnal Riset Ilmiah*, *2*(10), 4473–4481.

Ulya, I., Amilda, & Otoman. (2024). Timah dalam Pembentukkan Budaya dan Ekonomi Masyarakat Bangka. *Tanjak: Jurnal Sejarah Dan Peradaban Islam, 4*(1).

Winarso, G., Eko Kurniawan, dan, Pemanfaatan Penginderaan Jauh, P., & Tinggi Teknologi Angkatan Laut TNI-AL, S. (2014). Pemanfaatan Data Penginderaan Jauh Untuk Mendukung Perencanaan Operasi Keamanan Laut Di Laut Arafuru (Remote Sensing Application For Supporting Maritime Security Operations At Arafuru Sea). In *Jurnal Penginderaan Jauh* (Vol. 11, Number Desember).

Yanto, A., Salbilla, F., & Sitakar, R. (2023). Implikasi Resentralisasi Kewenangan Pertambangan Timah Terhadap Potensi Pendapatan Daerah di Bangka Belitung. *JURNAL INTERPRETASI HUKUM*, *4*(2). https://doi.org/10.55637/juinhum.4.2.7756.344-357

Zhao, Z., Qiu, S., Chen, F., Chen, Y., Qian, Y., Cui, H., Zhang, Y., Khoramshahi, E., & Qiu, Y. (2023). Vessel Detection with SDGSAT-1 Nighttime Light Images. *Remote Sensing*, *15*(17). https://doi.org/10.3390/rs15174354

Copyright © 2022 The Authors. Published by Department of Earth Science and Technology, Universitas Negeri Gorontalo
This work is licensed under a Creative Commons Attribution (CC-BY) 4.0 International License.