remote sensing

MDPI

Article

Automated Shoreline Segmentation in Satellite Imagery Using USV Measurements

Antoni Jaszcz [1,2,*], Marta Włodarczyk-Sielicka [3], Andrzej Stateczny [4], Dawid Połap [1] and Ilona Garczyńska [2,3]

[1] Faculty of Applied Mathematics, Silesian University of Technology, Kaszubska 23, 44-100 Gliwice, Poland; dawid.polap@polsl.pl

[2] Marine Technology Ltd., Roszczynialskiego 4/6, 81-521 Gdynia, Poland; i.garczynska@pm.szczecin.pl

[3] Department of Hydrography and Spatial Analysis, Faculty of Navigation, Maritime University of Szczecin, Wały Chrobrego 1-2, 70-500 Szczecin, Poland; m.wlodarczyk@pm.szczecin.pl

[4] Faculty of Navigation, Gdynia Maritime University, 81-87 Morska St., 81-225 Gdynia, Poland; a.stateczny@wn.umg.egu.pl

* Correspondence: aj303181@student.polsl.pl

Abstract: Generating aerial shoreline segmentation masks can be a daunting task, often requiring manual labeling or correction. This is further problematic because neural segmentation models require decent and abundant data for training, requiring even more manpower to automate the process. In this paper, we propose utilizing Unmanned Surface Vehicles (USVs) in an automated shoreline segmentation system on satellite imagery. The remotely controlled vessel first collects above- and underwater shoreline information using light detection and ranging (LiDAR) and multibeam echosounder (MBES) measuring instruments, resulting in a geo-referenced 3D point cloud. After cleaning and processing these data, the system integrates the projected map with an aerial image of the region. Based on the height values of the mapped points, the image is segmented. Finally, post-processing methods and the k-NN algorithm are introduced, resulting in a complete binary shoreline segmentation mask. The obtained data were used for training U-Net-type segmentation models with pre-trained backbones. The InceptionV3-based model achieved an accuracy of 96% and a dice coefficient score of 93%, demonstrating the effectiveness of the proposed system as a source of data acquisition for training deep neural networks.

Keywords: USV; LiDAR; multibeam echosounder; image; fusion; masks; segmentation; automatization

check for updates

Citation: Jaszcz, A.; Włodarczyk-Sielicka, M.; Stateczny, A.; Połap, D.; Garczyńska, I. Automated Shoreline Segmentation in Satellite Imagery Using USV Measurements. *Remote Sens.* **2024**, *16*, 4457. https://doi.org/10.3390/rs16234457

Academic Editor: Javier Marcello

Received: 11 September 2024
Revised: 16 November 2024
Accepted: 24 November 2024
Published: 27 November 2024

1. Introduction

The task of shoreline segmentation is critical for proper coastal management. Creating accurate coast outlines is necessary in many fields, such as environmental monitoring, land-use planning and coastal geomorphology. Due to the seafront variability and limitations (as restricted and difficult to reach areas) connected with conventional aerial mapping, the obtained land water can be imprecise. Furthermore, the once-examined shoreline can be changed significantly due to many processes, such as moving tides (caused by the gravitational force of the Moon and the Sun), erosion and human activity. These changes require manual adjustments to account for them. Therefore, there is an opportunity to use intelligent methods in coastal monitoring [1], as there is a demand for reliable and automated systems to facilitate the task of determining the coastline [2]. Accurate shoreline segmentation is crucial, especially for environmental monitoring because it serves as a foundation for understanding and managing dynamic coastal environments, particularly in the context of coastal erosion and other climate-driven changes. This is especially visible in the aftermath of the 2024 Central European floods, which caused great damage along the Danube and the Oder rivers.

Unmanned Surface Vehicles (USVs) are used practically to perform measurements, monitor the environment, and even analyze underwater and surface data [3]. Attention should be paid to equipping these vehicles with various sensors, including a thermometer for measuring temperature, a sonar for bottom analysis, radars, or a camera for observing the vehicle's surroundings [4]. A given vehicle may collect more real-time measurement data for subsequent or current analysis, depending on the installed sensors. For this purpose, wireless communication models with a ground station or server where the acquired data are stored are built. This action allows us to keep a copy of the data in the event of a vehicle failure. USVs present an innovative solution to the challenges of shoreline segmentation. The analysis of both the above- and underwater features of the coast provides precise information that complements aerial imagery.

A key role of USVs is coastal mapping, which can be defined as the process of collecting, analyzing and visualizing data related to the coastal zones. From a practical point of view, waterfront mapping allows for both topographic and structural monitoring, along with the proper shoreline erosion analysis [5]. Most often, the data necessary for segmentation are obtained, among others, using satellite images and spatial data [6–8]. As a base method, it is very efficient in terms of time and resources, allowing surveys to be carried out on large areas. This advantage, however, comes with the disadvantage of low detail of the data gathered. A complementary method is aerial photogrammetry, which considers the high resolution of measurement data and, therefore, high detail. It is essentially a half-measure, used to improve the accuracy of the created shoreline mappings locally. Unfortunately, due to the limitations connected with the natural layout of the shorelines, the aerial view can be easily obstructed. Furthermore, it requires a great amount of manual labor, as the surveys are performed locally and must traverse through large areas. Moreover, the images collected must be combined and processed by experts. This creates a major problem in terms of monitoring potential changes, as there is no easy way to verify whether the mapping is still valid or needs to be adjusted due to changes in the shoreline (shoreline corrosion, rising water levels, etc.). To do so, another aerial survey needs to be performed, resulting in high maintainability costs of this shoreline monitoring method. To facilitate these tasks and reduce the operational burden, USVs can be used to perform quick local surveys, which complement the aerial imagery of the terrain, resulting in more precise shoreline mapping [9].

To perform shoreline surveys, USVs must be equipped with the appropriate measuring instruments. The following sensors are sonar (such as MBES), LiDAR, i.e., [10]. These technologies enable the creation of sonar images or point clouds. Sonar enables underwater analysis, and LiDAR enables surface analysis. In addition to these sensors, other instruments increase the accuracy of surveys and provide more possibilities for further data processing and analysis. This analysis is primarily related to three tasks. The first one is geo-referencing, i.e., assigning exact coordinates of the acquired data. A practical application of geo-referencing is the ability to locate specific objects quickly. The second task is to model the waterfront area, which enables a comparison and control of ongoing environmental changes. The third task is to analyze the acquired data and the generated models to pay attention to coastal erosion [11] or vegetation changes. In a broader context, the above functions allow for environmental protection, increased safety and spatial planning.

The potential of waterfront analysis through mapping is vast, with numerous applications. However, it is crucial to consider the context of the processing methods [12,13]. Automating these processes often involves using advanced artificial intelligence methods that require substantial data for accurate model training [14]. An example of using artificial neural networks is the analysis of long-term coastal changes, which are shown in [15]. The work draws attention to the increase in sea level and waves, significantly affecting the environmental condition. To this end, the authors used convolutional neural networks to analyze and predict coastal retreat. Attention was paid to the possibilities of practical use, such as the possibility of supporting coastal surveillance and management. Another interesting analysis approach is the detection of shorelines to easily extract the shore [16]. A different approach is to use two collaborative neural network models to continuously estimate the waterline image map [17]. The mentioned research focuses on data extraction from images obtained from the USV level, but research is also carried out using SAR satellite images [18]. Different approaches were also analyzed in [19], where different methods were investigated, like the thresholding technique, k-means, random forest, etc. This research also uses post-processing tools like image binarization. It is worth discussing different approaches to find the best methods for different data.

A significant problem is USV control, which often depends on weather conditions. Solutions based on autonomous navigation functions are modeled, an example of which is cooperative navigation [20]. As part of the research, a method was built to estimate traffic information based on the extended Kalman version, using data obtained from various sensors. It is worth paying attention to geometric information and the ideas of the leader and the leader follower. In this model, the leader returned target points for the others in the context of information exchange. Route planning methods are also based on specific environmental conditions and maps. An example is a method that uses nautical maps and historical data on sea currents [21]. This solution recognizes the need to optimize various factors to guarantee safety, energy savings and efficiency. Another navigation algorithm is based on additional mooring and unmooring algorithms for analyzing various obstacles along the route.

Autonomous vehicles also require reliable self-navigation methods to traverse different, often rapidly changing terrain. One of the approaches is to use stereoscopic vision [22], where the data are analyzed to detect obstacles' location and dimensions. Recursive estimation techniques were chosen to enable this, making it possible to obtain a real-time fusion of camera and navigation data. Recurrent methods are also extended to neural networks, as shown in [23]. Scientists paid attention to forecasting time series of ship trajectories using models built with LSTM layers, and the selection of hyperparameters was made using an optimization algorithm.

Based on the literature analysis, it was noticed that mapping the waterfront can be performed by using satellite photos or photos obtained from drones or planes. However, accurate waterfront extraction involves the analysis of various objects, which rely on one's perception. For this purpose, we propose to create a framework that will fuse data from multiple sensors to create automatic masks for mapping the seafront. The shoreline data can be collected by a USV, gathering both underwater and shore information through a multibeam echosounder and LiDAR adequately. By combining these data, a 3D point cloud can be modeled, which creates a visualization of the shore as a whole. We propose to further utilize this visualization by combining it with satellite images, which will enable the creation of much more accurate masks of the waterfront. The masks generated this way can be used in training segmentation networks for delineating shorelines. The practical possibilities of such fusion are essential when obtaining large amounts of data are complex and the proposed technique can automate this otherwise mundane and resource-intense process. This study aims to develop a robust automated shoreline segmentation system that integrates diverse data acquisition methods—USV, LiDAR, MBES, and satellite imagery—to create a unified and accurate shoreline detection framework. This combination of technologies offers a novel approach to shoreline segmentation by uniting above-water, underwater, and aerial perspectives into a comprehensive and reliable model. The main contributions of this paper are as follows:

- Development of an automated shoreline segmentation system using local measurements collected by USV;
- A method of integrating LiDAR and MBES data with satellite imagery to create segmentation masks;
- Post-processing pipeline for partially segmented shoreline images;
- Performance analysis of the pre-trained encoders used in the U-Net model for shoreline segmentation tasks.

2. Methodology

2.1. LiDAR and MBES

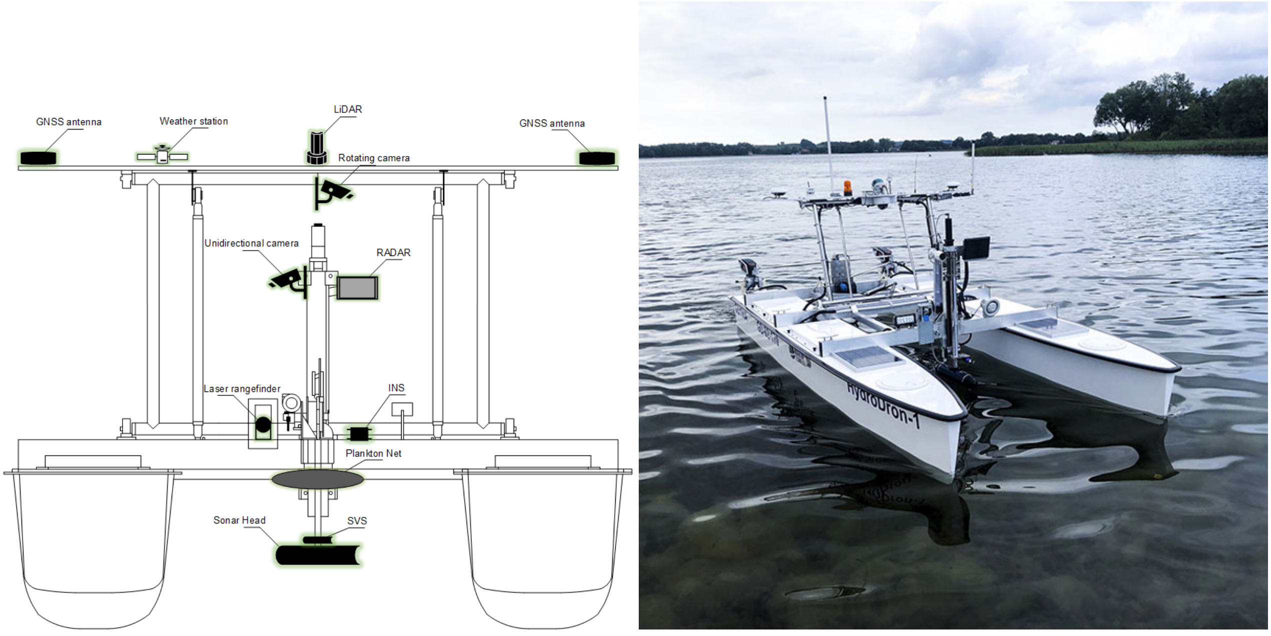
Laser imaging, detection, and ranging (LiDAR) is one of the sensing methods that utilizes a series of pulsed lasers to map objects based on the retrieved distance. Many remote systems utilize this technology to create high-precision 3D maps of terrain. Although it is most commonly used in aerial mapping, it can also be used in vertical mapping, such as generating detailed topographic information about shorelines.

A multi-beam echosounder (MBES), on the other hand, involves multiple sound beams to acquire information about the terrain. It is widely used for seafloor mapping and other underwater imaging tasks. By emitting fan-shaped sound waves from the transducer, the return time of these waves (which are reflected by an object upon collision) is measured and used to calculate the depth of terrain.

By utilizing both LiDAR and MBES technologies, the proposed system can extract information on underwater and land terrain, crucial for designating precise shorelines.

2.2. Unmanned Surface Vehicle

The presented research uses a USV named HydroDron [24]. It is a mobile unit built by Marine Technology, Poland, and dedicated to performing measurement missions in restricted waters. The vehicle moves autonomously in a given area, collecting data using built-in navigation sensors on the mast. In addition, the vehicle contacts a ground station, where the collected data are saved via a wireless network. Equipped with many sensors and navigational devices, it can navigate waters autonomously while performing surveys. The measuring devices include a LiDAR Velodyne VLP-16 sensor and a PING 3DSS-DX-450 echo sounder. This equipment, along with navigational devices (depicted in Figure 1a), creates an autonomous vessel that can conduct surveys in diverse environments. Our research used the HydroDrone watercraft to gather LiDAR and MBES data from various shorelines.



(**a**) (**b**)

Figure 1. The HydroDrone used for shoreline data acquisition. This USV is designed for measurement missions in confined or restricted waters. It can conduct a range of measurements, including bathymetric, sonar, LiDAR, and others. The vehicle operates autonomously, following a pre-set path, but can also be controlled remotely. (**a**) Scheme of the HydroDrone USV and its measuring and navigation instruments; (**b**) Picture of the HydroDrone USV taken during a survey mission.

2.3. Proposed System for Shoreline Segmentation

The overview of the proposed system is displayed in Figure 2. The presented methods can be divided into two stages. In the first stage, a USV equipped with LiDAR and MBES is sent to the area of interest, gathering necessary surface and subsurface data. Then, the collected data are cleaned. That involves removing the existing outliers and thinning the data points (as some of the regions of the gathered 3D point cloud are too dense). On the contrary, if some parts of the 3D model are too sparsely populated, the interpolation algorithms are involved to produce smooth results. Depending on the required precision, 3D modeling with different mesh densities can be used. For shoreline segmentation, a density of 0.5 m is sufficient. As the end result of the stage, a 3D point cloud representing the shore is acquired. The 'x' and 'y' coordinates represent the UTM system, while the 'z' coordinate is in meters. In the second stage, we obtain a segmentation mask using the obtained point cloud. First, the points are projected onto the 2D plane determined by the x and y coordinates. While doing so, the points are classified based on their height, using the θ threshold parameter. This parameter indicates the subjective water level, which can be impacted by the height of the LiDAR device and other factors. This whole step essentially creates a bird-eye view of the land and water. Further processing requires determining the boundary of the point cloud and translating the obtained limitation values to the geographic coordinate system (GCS). Next, having the center point of the projection, an aerial image of the region is acquired using third-party API, such as Google's Earth Engine. While using aerial imagery, its resolution needs to be considered. Typically, it is measured with ground sample distance (GSD), as shown in Equation (1). It is calculated using sensor size, focal length (both in millimeters), image size (pixels) and altitude of the capturing device (meters). The GSD measures the smallest perceivable distance in the image (what distance does one pixel represent in reality). Depending on the satellite the data are acquired from, GSD may differ. Landsat satellite series (4, 5, 7 and 8) provide GSD of 30 m for most bands. On the other hand, Sentinel-2 (maintained by the European Space Agency) provides imagery with SGD = 10 meters for the RGB bands. Knowing the SGD of the satellite image, we can then match the points from the projection to pixels in the photo layout and crop the image to the considered selection (designated by the projection). This results in a semi-completed binary mask, as some of the pixels are unlabeled. Based on the known points, we can utilize a simple classifier (like k-NN) to match the rest of the points based on the pixel coordinates and RGB values from the satellite image. This results in the completed mask of the shore being created automatically. Lastly, post-processing is performed, removing the misclassified pixels and smoothing the transition between land and water in the segmented shoreline. The following subsections describe both stages in more detail.

$$GSD = \frac{sensor\_size \times altitude}{focal\_length \times image\_size} \tag{1}$$

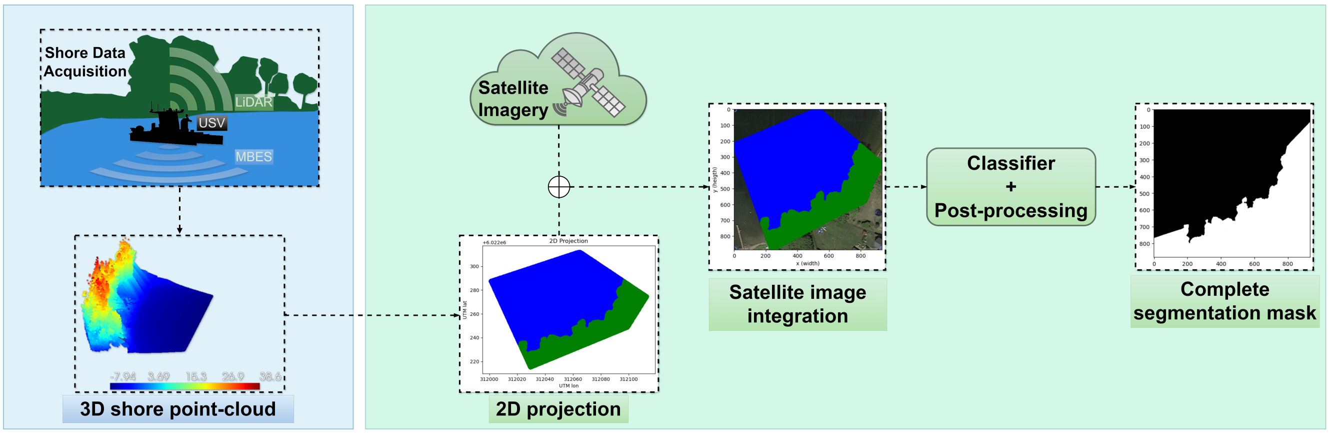


Figure 2. Framework of the proposed system. The blue section represents the data acquisition and preparation stage. In this stage, data are gathered via USV and processed for further usage. The green segment presents an integration of the obtained 3D model of the shore with the satellite imagery, resulting in a complete segmentation mask of the waterfront.

2.3.1. Data Preparation

In order to collect data, a USV vehicle (like the one described in Section 2.2) can be used. Such vessel collects above-water LiDAR data and underwater MBES data using the equipped sensors. As the USV navigates along the shoreline, the LiDAR emits laser pulses that reflect off the shore, capturing detailed elevation and surface data above the waterline. Simultaneously, the MBES emits acoustic beams that bounce off the seabed, recording the depth and contour of the underwater portion of the shoreline. The data from both sensors are time-synchronized and geo-referenced using navigational devices, such as GNSS antennas and the onboard computer. This ensures accurate alignment of the collected 3D mesh. By combining the acquired data and clean-up process, a comprehensive 3D point cloud representing both the above-water and underwater features of the shoreline is obtained. This results in a detailed and geo-referenced model that can be used further in the proposed system.

2.3.2. Processing Module

To better demonstrate the second stage of the system, the exact steps are described in Algorithm 1. The θ threshold highly depends on the data acquisition method used in the proposed system. In the proposed approach, θ parameters were determined based on the height of the LiDAR device on the USV and its total draught. The integration of the projected points and satellite image is based on the translated min/max coordinates and the dimensions of the image. To minimize the computational requirements, it is proposed that every point is assigned to a pixel based on its UTM values without translation to GSC. This can be carried out for every point using the following formula (Equation (2)). The resulting x and y values are the coordinates in the satellite image layout.

$$
\begin{aligned}
x &= \frac{lon - minLon}{maxLon - minLon} \times width \\
y &= \frac{maxLat - lat}{maxLat - minLat} \times height
\end{aligned}
\tag{2}
$$

Having mapped the points to the image, a simple classifier can be used to predict the unassigned pixels. We propose the k-NN algorithm in the presented system, as it works as a simple yet very effective segmentation algorithm in the considered scenario. This is because the quantity of data for a single image is not extensive and the known labels are already densely mapped. The algorithm considers RGB values from the satellite image and x, y coordinates, totaling five features (whose values are normalized to the interval [0, 1]). This provides the algorithm with both quality and positional information about the surrounding pixels. A simple post-processing is performed to improve the quality of the generated masks further. First, to address the issue of noise at the water–land border at individual land pixels misclassified as water, the majority filter is applied to the mask. Furthermore, opening and closing morphological operations are introduced to the mask to prevent larger misclassified areas. The opening operation performs erosion and dilation, while closing does the opposite. Erosion shrinks in a binary image, while dilation thickens objects in a binary image. Given an image X and structuring element (kernel) K, as well as a translation of K by the vector Z denoted as $(X)_z$, erosion can be presented as in Equation (3) and dilation as in Equation (4).

$$X \ominus K = \{z \mid (K)_z \subseteq X\} \tag{3}$$

$$X \oplus K = \{z \mid (K)_z \cap X \neq \emptyset\} \tag{4}$$

Both majority filter and morphological information require kernel size k. The influence of this parameter is analyzed further in this paper. After post-processing, the completed and cleaned segmentation masks of the shore are returned.

Algorithm 1: Complete and clean mask using k-NN and morphological operations

```
Concatenate LiDAR and BMES data into data
Set $elevation\_threshold = 0$
Extract water points (elevation $\leq$ elevation_threshold)
Extract land points: (elevation $>$ elevation_threshold)
Find minimum and maximum x and y coordinates:
  MIN_X, MAX_X, MIN_Y, MAX_Y
Define UTM and WGS84 coordinate systems:
  utm_proj, wgs84_proj
Convert UTM to geographical coordinates:
  minLon, minLat, maxLon, maxLat
Calculate center latitude and longitude:
  center_lat, center_lon
Get satellite image: image, based on the coordinates
Get image dimensions: width, height
Convert land and water UTM coordinates to pixel coordinates:
  land_pixel_coordinates, water_pixel_coordinates
Create an empty mask of size (height, width)
Assign labels to the mask at the specified points: 1 for land, 0 for water
Extract labeled points: labeled_points, and their labels: labels from mask
Extract features (coordinates and pixel values) from image at labeled_points
Normalize features for labeled points
Train k-NN classifier with features and labels
Extract unlabeled points: unlabeled_points from mask
Extract features for unlabeled points from image
Normalize features for unlabeled points
Predict labels for unlabeled_features using k-NN classifier
Create the completed mask by assigning predicted labels to unlabeled_points in mask
Apply majority filter to clean the completed mask: cleaned_mask
Apply morphological operations (opening and closing) to further refine cleaned_mask
Return the cleaned mask: knn_mask.png
```

3. Experiments

3.1. Collected Data

The shoreline data were collected from eight different areas located in northwestern Poland during the surveys conducted between 2022 and 2024. Figure 3 shows the areas related to data acquisition. Survey campaigns were planned along the waterfronts in Gdynia, covering locations such as Marina Yacht Park, Breakwater, and the Pomorskie and Słowackie Quays (Example Figure 4b). In addition, data were also collected in more secluded locations (Lake Kłodno, Zawory). Figure 4a presents the location of the two different survey campaigns. Each location presented a unique set of challenges for shoreline recognition tasks. Crowns of trees growing along the shore cover visible ground, making it difficult to delineate the shoreline. On the other hand, in the more attended areas, multiple ships are docked along the concrete waterfronts. This can especially impede the segmentation task, as the ability to retake the aerial image (if at all possible) is severely limited, and such a situation usually requires human interaction to remove the vessel from the satellite image manually. However, with the help of surface shoreline mapping, the missing references can be collected when the ships depart. This can be observed in Figure 4c, where the proposed system was able to ignore the vessels visible in the aerial images successfully.

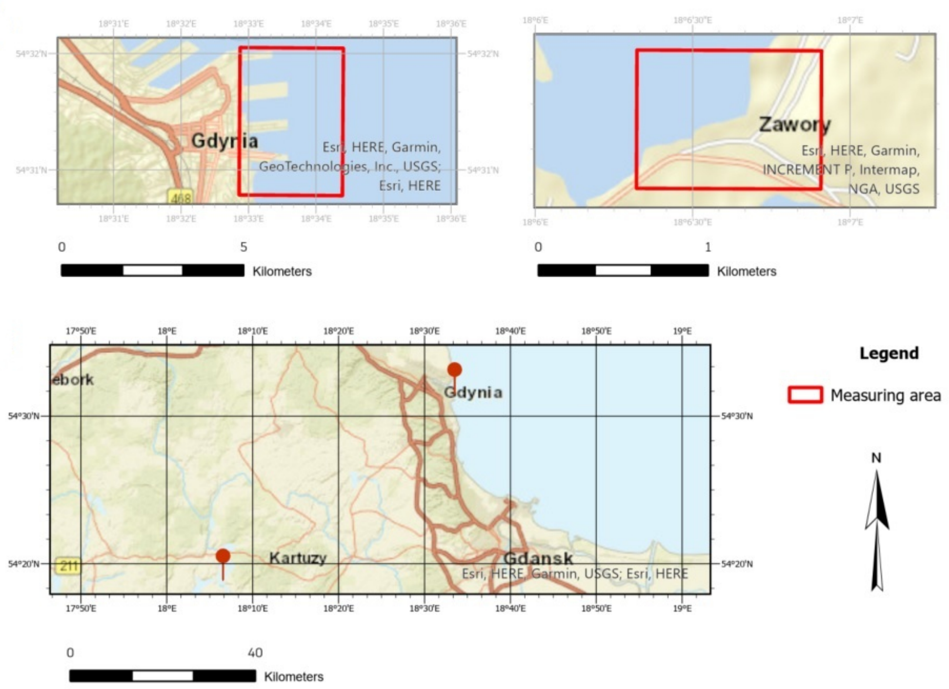
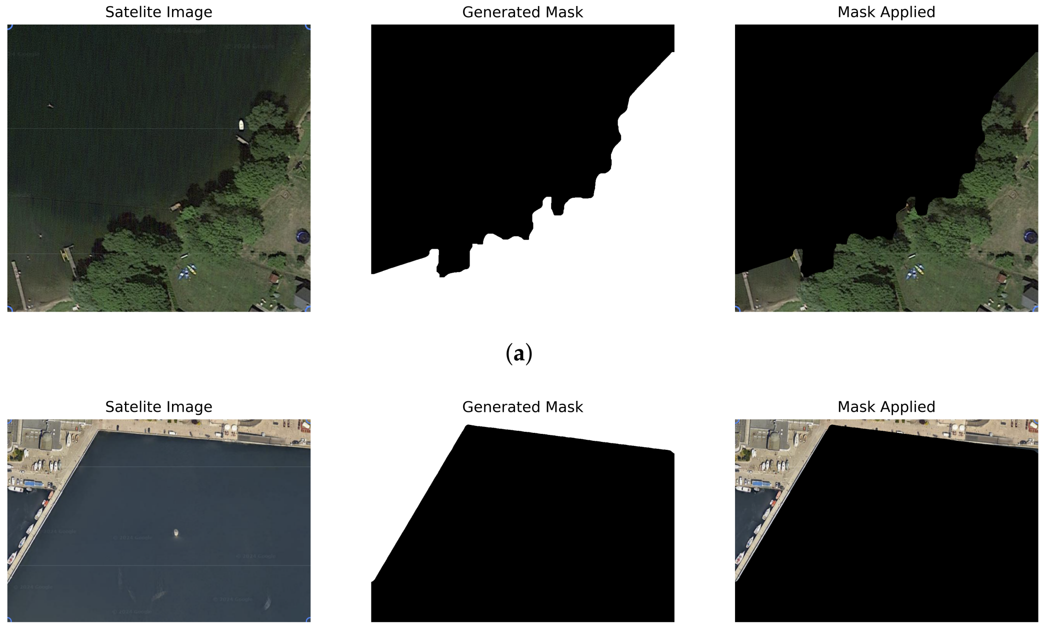

Gdynia
Esri, HERE, Garmin, GeoTechnologies, Inc., USGS; Esri, HERE
0
5
Kilometers
Zawory
Esri, HERE, Garmin, INCREMENT P, Intermap, NGA, USGS
0
1
Kilometers
Gdynia
Kartuzy
Gdansk
Esri, HERE, Garmin, USGS; Esri, HERE
0
40
Kilometers
Legend
Measuring area
N

Figure 3. Areas where USV surveys were conducted.

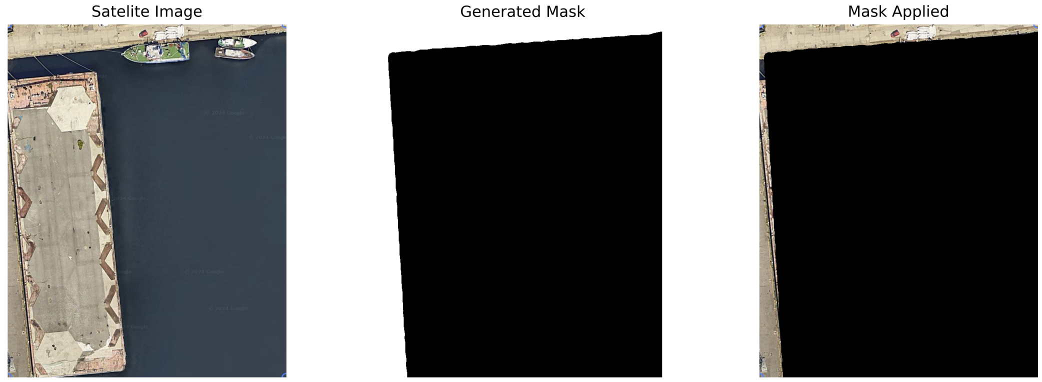
During data acquisition, information was recorded in real time on onboard computers mounted on the autonomous vessel. LiDAR data registration was conducted simultaneously with bathymetric measurements. Hypack 2024 software was used for both data recording and processing. This advanced hydrographic software provides a suite of tools that facilitate bathymetric measurements and LiDAR data registration. It enables the planning of data acquisition by creating designed survey lines, filtering data in both automatic and manual modes, as well as integrating data from various sensors. The data were stored in HSX and RAW file formats.



(**a**)

(**b**)

Figure 4. *Cont.*



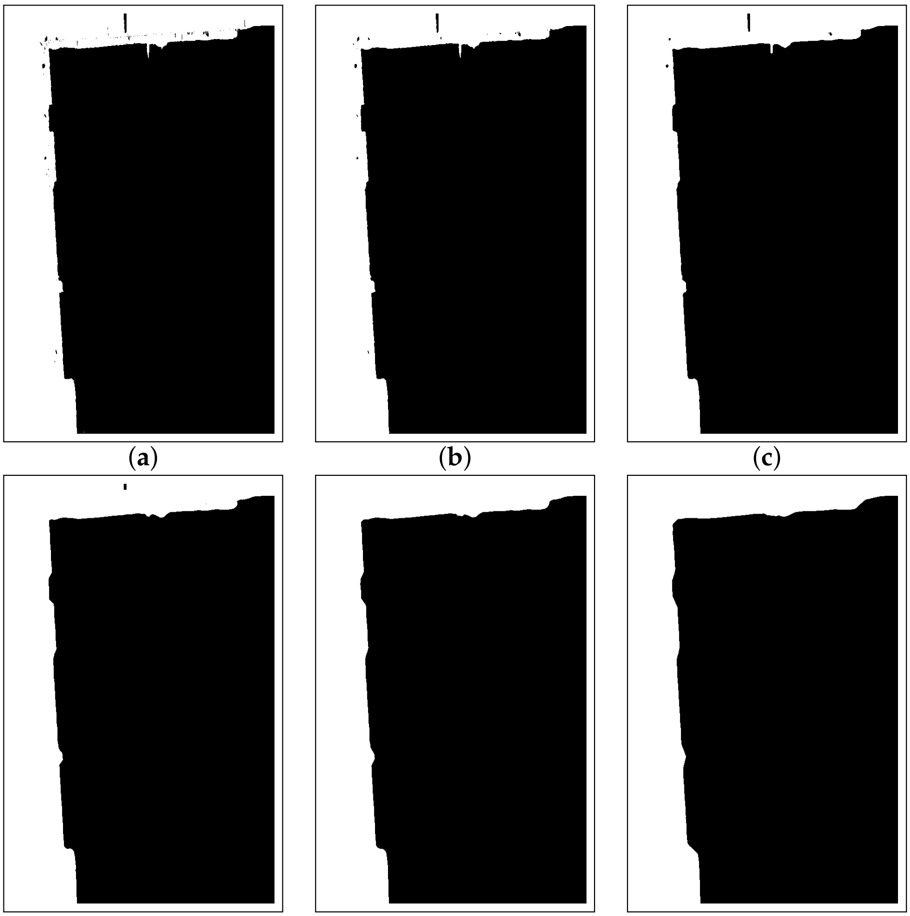
(**c**)

Figure 4. Examples of different waterfronts included in the experiment along with the predicted and segmentation masks. Each of the presented examples poses a unique challenge for the segmentation process. (**a**) Shoreline with dense vegetation. (**b**) Concrete quay. (**c**) Vessels docked long-side to the wharf.

3.2. Obtained Masks

As previously described in Section 2.3.2, the kernel's size in post-processing significantly impacts the end result. To better demonstrate this, Figure 5 presents the example mask without any post-processing, as well as the masks resulting from different kernel sizes. The larger the size of the kernel, the larger the misclassified holes are patched. However, a greater K parameter also produces more oval-shaped edges, which is not desirable in excess. Hence, based on the conducted experiments, a 15 × 15 kernel was chosen for the post-processing methods as a compromise.

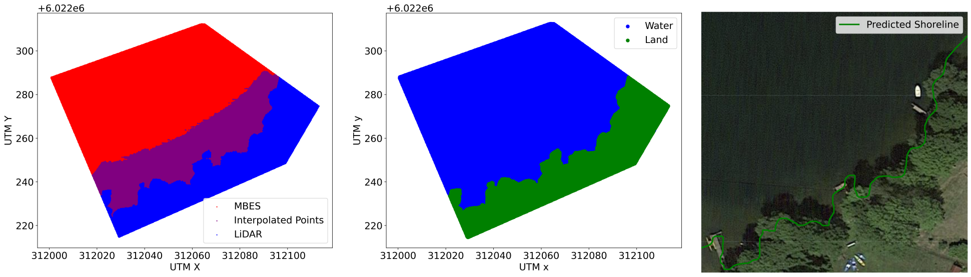
Figure 4 shows three examples of predicted shorelines. It is worth noting that each presents a different segmentation challenge. A wooded shoreline is presented in Figure 4a. Blending colors and shadows cast by trees are problematic for the computer-vision-based segmentation models. The results of the various steps of our method for this sample are presented in Figure 6. However, the proposed method performed well and, with the local USV data, was able to create the mask reliably. However, due to the missing surface data, the satellite image shows that parts of the shoreline on the left and right were not covered. Without this reference, the system completes the shoreline with a straight line. A similar limitation can be observed in Figure 4b. Even though the completed mask perfectly resembles the shoreline's shape perfectly, the water visible on the left side of the pier was also classified as land. The surface data are insufficient to fully segment the obtained image, as the USV performed the measurements on the right side of the pier. Despite this, such an aspect of the system can also be beneficial. In Figure 4c, the visible boats and rafts were absent during the USV survey when the satellite image captured them. Because of that, the presented system could cleanly capture the true shoreline without boats.



(**a**) (**b**) (**c**)

(**d**) (**e**) (**f**)

Figure 5. Effect of different kernel sizes in the mask post-processing algorithm. The greater the value of k, the smoother the result. (**a**) No post-processing performed; (**b**) $k = 3$; (**c**) $k = 7$; (**d**) $k = 11$; (**e**) $k = 15$; (**f**) $k = 25$.



(**a**) (**b**) (**c**)

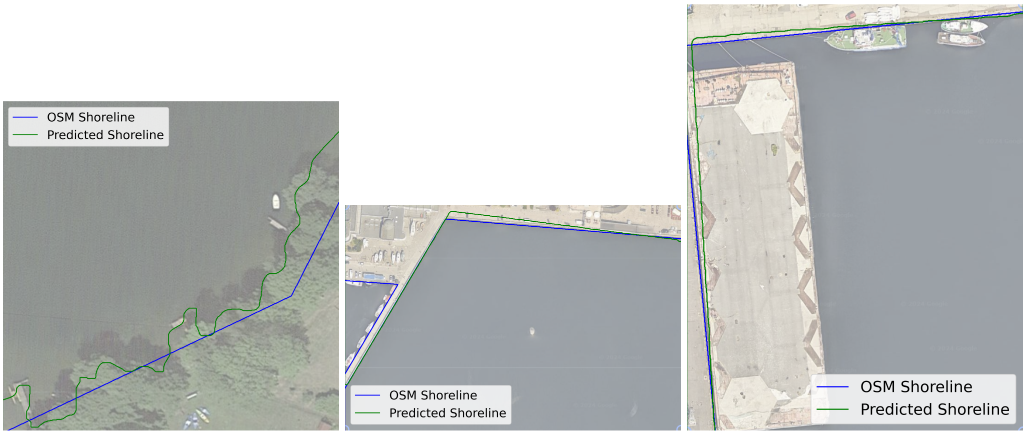
Figure 6. Resultant projections of each step of the proposed method shown in one of the experimental areas. (**a**) Combined shoreline projection. (**b**) Complete shoreline projection. (**c**) Continuous shoreline projection.

3.3. Shoreline Comparison

To extend the research, an additional comparison between the extracted shoreline vectors obtained through our proposed method and the shoreline data provided by the OSM platform was performed. The visual comparison is presented in Figure 7, with corresponding Mean Perpendicular Distance (MPD) values obtained from the samples. The MPD is calculated by averaging the shortest distances from points on a predicted line (or polygon) to the corresponding points on the true line. The formula is as follows:

$$MPD = \frac{1}{N}\sum_{i=1}^{N} d_i, \tag{5}$$

where N is the total number of anchoring points on the predicted line, and d_i is the perpendicular distance from the ith point on the predicted line to the nearest point on the true line.



(**a**) (**b**) (**c**)

Figure 7. Comparison between shorelines obtained from the segmentation masks and the vectorized OSM. (**a**) $MPD = 39.75$; (**b**) $MPD = 9.07$; (**c**) $MPD = 5.32$.

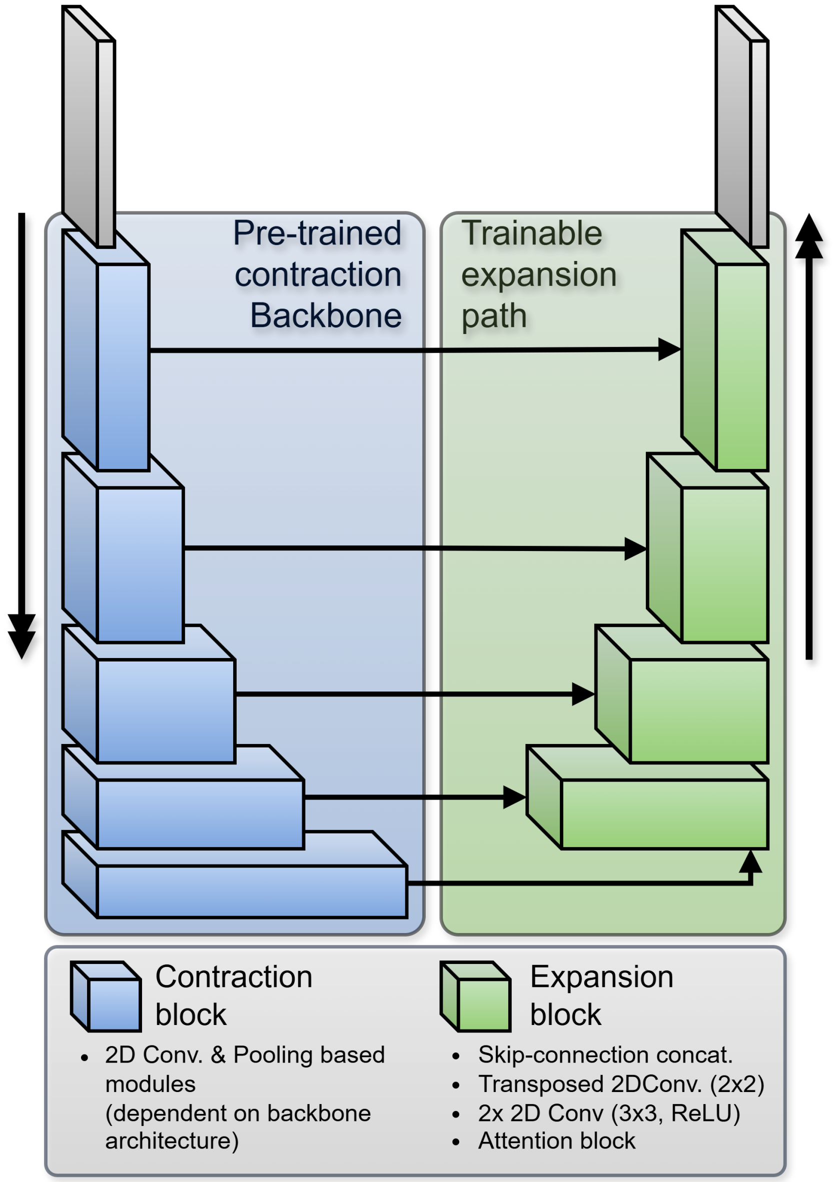
Based on these results, the following conclusion can be drawn. First, there is a difference in shorelines. This is mainly due to the different precision available through the proposed method and the shoreline outlining using satellite imagery exclusively. Technologies like LiDAR (light detection and ranging) and MBESa (multibeam echosoundera) provide dense, high-resolution 3D point clouds, making them ideal for capturing fine-scale coastal features, such as small inlets or cliff edges, while satellite imagery generally has lower spatial resolution (especially in freely available data sources like Landsat (30 m) or Sentinel-2 (10 m). As the shoreline obtained from the satellite imagery is based on the image, it is strongly align, while the proposed shoreline—obtained from local measurements—deviates from that. This is caused by both the GPS inaccuracy and the expected satellite imaging error. Furthermore, we can see that the highly vegetated waterfront was simplified in the OSM, while the proposed method designated a shoreline taking into account dense trees along the shore (Figure 7a). However, the proposed method suffers from representational inaccuracy, if the survey does not take into account the whole shoreline. As can be observed in Figure 7b, part of the port basin located on the left side of the image was not included in the shoreline, since the USV did not perform the measurements on the opposite side of the peer. Nevertheless, the MPD values validate the correctness of the designated shorelines, which provide more precise representation of the waterfront, than the simplified version obtained solely from the aerial imagery.

3.4. Training Segmentation Models

To validate the obtained masks as ground-truth data for neural networks in segmentation tasks, we conducted experiments using U-Net-type architecture networks with different pre-trained encoder backbones. The architecture of the model in detail is presented in Figure 8. For the experiments, we chose five different models pre-trained on the ImageNet dataset:

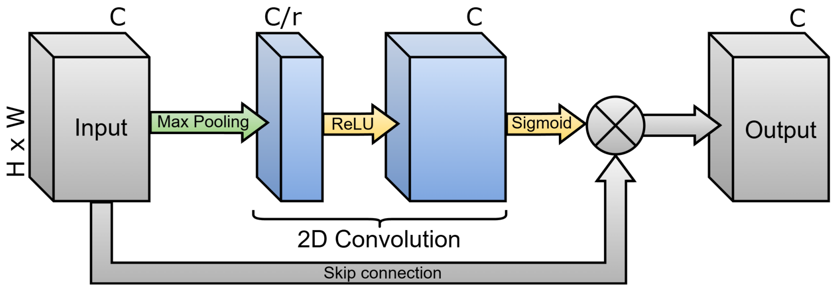
- VGG16 [25]—a widely-known CNN architecture known for its simplicity and depth. It comprises only 16 layers (13 of which are convolutional layers).
- ResNet50 [26]—deep network with residual blocks, which provide better information flow within the network. It is used for a variety of tasks that include feature extraction.
- MobileNetV2 [27]—a lightweight model designed to run on edge devices. It contains smaller residual blocks. It is computationally efficient and, despite its size, often very effective.
- InceptionV3 [28]—a deep network that utilizes a combination of convolutional paths across consecutive layers to capture different aspects of the image at various scales.
- EfficientNetB0 [29]—the smallest model in the EfficientNet family with high scaling across the input dimensions. Because of this strategy, it is computationally efficient.

Each image and segmentation masks were resized to 1024 × 1024 and divided into 256 × 256 patches, which were then fed onto the network. Random flips and rotations were applied as augmentation methods. Each model was trained for 200 epochs with the ADAM optimizer, binary cross-entropy loss function, and 80:20 train/validation split.


Pre-trained contraction Backbone
Trainable expansion path
Contraction block
- 2D Conv. & Pooling based modules (dependent on backbone architecture)
Expansion block
- Skip-connection concat.
- Transposed 2DConv. (2x2)
- 2x 2D Conv (3x3, ReLU)
- Attention block

(**a**)

Figure 8. *Cont.*



(**b**)

Figure 8. Visual representation of the U-Net type network model used in the experiments on the generated segmentation masks. (**a**) Framework of the U-Net type segmentation model used in the experiments. The contraction path is built with a pre-trained convolutional backbone, while the expansion path consists of transposed convolutions and attention modules. (**b**) Scheme of an attention module used in the experiments. H, W, C represent the dimensions of the input data.

3.4.1. Evaluation Metrics

To evaluate the effectiveness of the segmentation models, two statistical measures of similarity between sets were introduced. The first is the Dice coefficient (Sørensen–Dice index). It is equal to the doubled number of common elements in the set of true values A and the predicted values B over the sum of their cardinalities. This is presented in the following Equation (6), along with the interpretation of the confusion matrix values.

$$Dice = \frac{2TP}{FN + FP + 2TP} = \frac{|A \bigcap B| \times 2}{|A| + |B|} \tag{6}$$

The second metric is Intersection Over Union (IoU), which is equal to the intersection of sets A and B divided by their union (7).

$$IoU = \frac{TP}{FN + FP + TP} = \frac{|A \bigcap B|}{|A \bigcup B|} \tag{7}$$

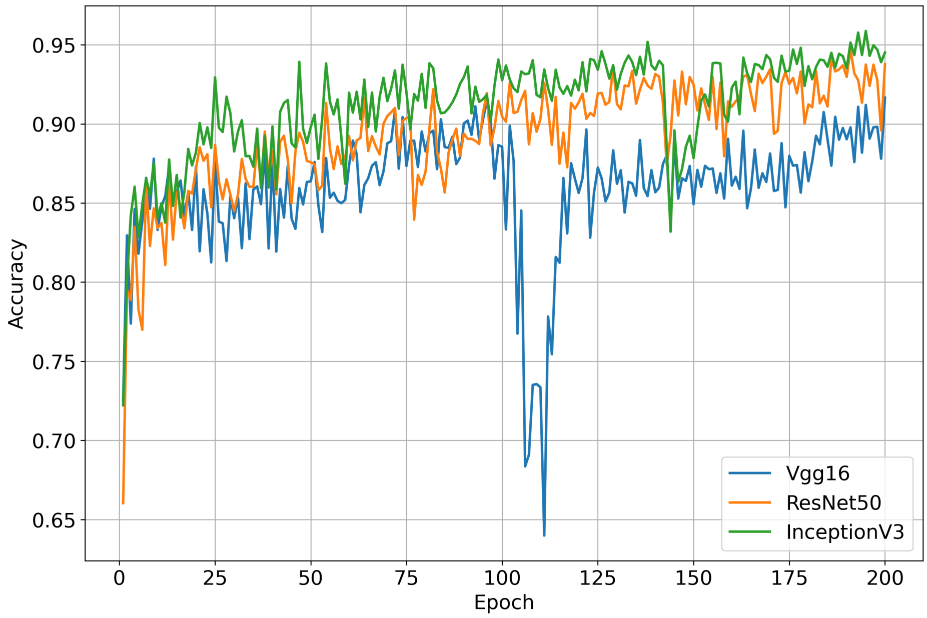
Despite the similarity of these measures, they can capture different flaws in the model's predictions. In terms of semantic segmentation, errors near the edges of the predicted object lower the score more than in the Dice coefficient (by the fact that TP predictions have a lower weight in the denominator). On the other hand, the Dice coefficient can be perceived as double the F1 score, providing valuable insight into the precision-to-recall ratio.

3.4.2. Training Results

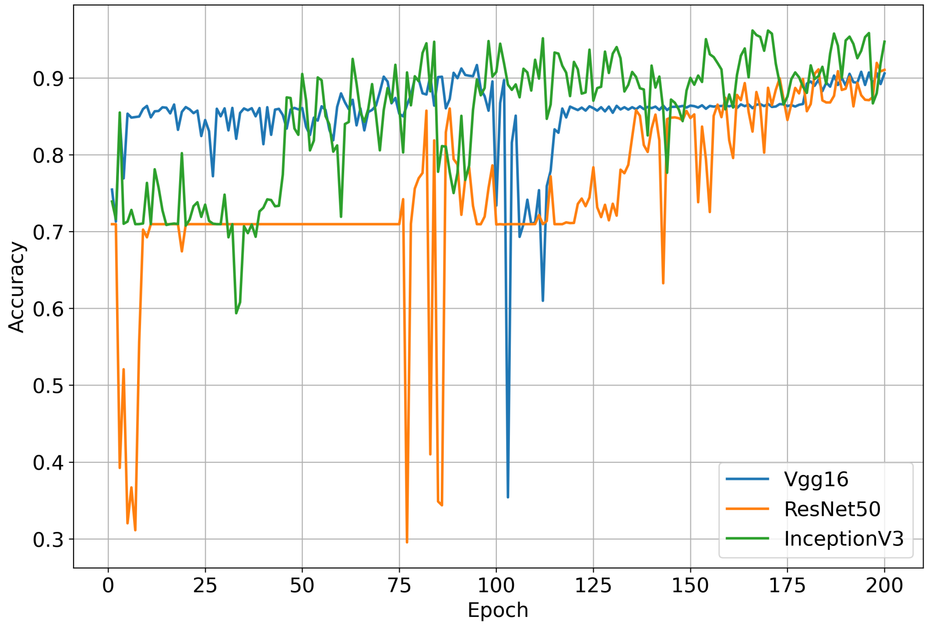
In the experiments we conducted, we measured the effectiveness of the model's predictions on the test set based on accuracy, Dice score, and IoU. The accuracy of the training and validation sets during training denser models is shown in Figure 9. The evaluation results are presented in Table 1. The total number of parameters in the model is also displayed. As can be seen, more lightweight backbones did not perform as well as more complex models. Both MobileNetV2 and EfficientNetB0 did not achieve an accuracy of above 90%, and the resulting IoU scores were below 70%. This indicates that such models tend to have problems with under- and over-classification along the edges, which is particularly undesirable in a given shoreline detection task. Thanks to the deeper but rather simple VGG16 backbone, the model was able to achieve decent results, compared to the nearly three times larger ResNet50 model. However, the model based on InceptionV3 performed best of all, with 96% accuracy and a 93% Dice score. This is due to its ability to capture relevant information at different scales.

Table 1. Comparison of the U-Net type models with different encoding backbones.

Backbone	Params	Accuracy (%)	Precision (%)	Dice (%)	IoU (%)
VGG16	25 M	90.58	84.64	83.57	71.78
ResNet50	72 M	91.05	89.53	83.57	71.78
MobileNetV2	5 M	86.47	69.02	80.62	67.54
InceptionV3	64 M	96.17	95.92	93.21	87.29
EfficientNetB0	7 M	78.81	60.88	72.34	56.67



(**a**)



(**b**)

Figure 9. Training process of the U-Net type models with different pre-trained backbones. Change in accuracy in the training and validation sets is presented. (**a**) Training accuracy during consecutive epochs during training; (**b**) Validation accuracy during consecutive epochs during training.

4. Discussion

The proposed shoreline segmentation system, which integrates UAV-based data collection with advanced image processing techniques, has demonstrated significant effectiveness in mapping shoreline boundaries with high precision. This method's primary advantage over conventional satellite or aerial imagery lies in its ability to collect localized, detailed information through the use of LiDAR and MBES technologies, able to obtain high-resolution 3D models of a waterfront. This research aimed to autonomously create segmentation maps of the shore supplementing the aerial imagery, which then could be used to train segmentation neural networks. The data obtained through the proposed method were then used to train U-Net type models with pre-trained backbones and attention modules.

The models achieved high accuracies with the InceptionV3 backbone, resulting in the best performance. This is mainly due to the recursive nature of the Inception-type network, working well with wind segmentation masks.

However, the approach also faces certain limitations. While UAV and USV-mounted LiDAR offer precise data collection capabilities, their effectiveness can be compromised in areas where objects like ships or piers obstruct the view of a shoreline. These obstructions can interfere with LiDAR data quality, introducing noise and reducing the accuracy of shoreline segmentation in those localized regions. Potentially, these problems might be addressed via interpolation algorithms. Furthermore, the volume of data collected in a single survey may still be limited compared to the expansive coverage of satellite imagery, making it challenging to generalize results over longer shorelines or diverse coastal environments without conducting multiple surveys. Thus, the proposed approach is rather limited and not easily scalable to broader shoreline mapping. Nevertheless, it serves as an effective method for obtaining local measurements, which are then used for creating segmentation masks of the waterfront without the need for manual annotation, which was the main goal of this study.

5. Conclusions

In this study, we developed an automated shoreline segmentation system that combines USV-collected local measurements with satellite imagery, designed specifically to address the challenges of precision and automation in coastal mapping. The system demonstrated high segmentation accuracy, creating geo-referenced 3D point clouds from LiDAR and MBES data, which were then used to create detailed binary segmentation masks of the shore. By employing k-NN classification and morphological operations, we enhanced the accuracy and smoothness of shoreline segmentation, yielding a reliable dataset suitable for training neural network models. Testing the method with a U-Net type model architecture, our results were particularly promising, with the InceptionV3-based model achieving 96% accuracy and a Dice score of 93%, highlighting the method’s potential in automated shoreline data acquisition for training segmentation models.

In terms of performance, our approach offers clear advantages over traditional methods by leveraging localized high-resolution data, which significantly improves segmentation quality. However, the method also has limitations, particularly in environments with obstacles near the shoreline, such as trees and vessels, which can create noise in LiDAR data and obscure portions of the shore. Furthermore, the computational resources required for processing and analyzing large LiDAR and MBES datasets can become substantial and hard to manage when scaling to larger regions or more diverse coastal environments. Despite that, the proposed approach serves as a great framework for detailed shoreline segmentation in important waterfront areas, especially with restricted accessibility. In the context of existing methods, the proposed framework allows for the analysis of a coastline performed using a combination of two types of data, which allows for increasing the accuracy of the obtained segmentation results.

To address these limitations, our future work will focus on optimizing our system to handle obstructions more effectively. Additionally, we plan to conduct surveys across varied locations with differing environmental and shoreline characteristics, which will allow us to assess the adaptability and feasibility of the approach on a broader scale. Important aspects also include temporal analysis, i.e., comparing the same areas regarding the data collected at a certain time interval. These improvements aim to make the system more robust and scalable, contributing to a comprehensive, automated solution for shoreline monitoring and coastal management applications.

Author Contributions: Methodology, A.J.; Softwar, A.J.; Validation, A.J.; Writing—original draft, A.J. and D.P.; Formal analysis, M.W.-S. and A.S.; Data curation, M.W.-S.; Resources, A.S.; Supervision, D.P.; Validation, I.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Centre for Research and Development 293 (NCBR) of Poland, under grant numbers LIDER/4/0026/L-12/20/NCBR/2021.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy considerations.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Prokop, K.; Połap, K.; Włodarczyk-Sielicka, M.; Jaszcz, A. End-to-end system for monitoring the state of rivers using a drone. *Front. Environ. Sci.* **2023**, *11*, 1303067. [CrossRef]
2. Sun, W.; Chen, C.; Liu, W.; Yang, G.; Meng, X.; Wang, L.; Ren, K. Coastline extraction using remote sensing: A review. *GISci. Remote Sens.* **2023**, *60*, 2243671. [CrossRef]
3. Kum, B.C.; Shin, D.H.; Lee, J.H.; Moh, T.; Jang, S.; Lee, S.Y.; Cho, J.H. Monitoring applications for multifunctional unmanned surface vehicles in marine coastal environments. *J. Coast. Res.* **2018**, 1381–1385. [CrossRef]
4. Sotelo-Torres, F.; Alvarez, L.V.; Roberts, R.C. An Unmanned Surface Vehicle (USV): Development of an Autonomous Boat with a Sensor Integration System for Bathymetric Surveys. *Sensors* **2023**, *23*, 4420. [CrossRef]
5. González-Teruel, J.D.; Torres-Sánchez, R.; Blaya-Ros, P.J.; Toledo-Moreo, A.B.; Jiménez-Buendía, M.; Soto-Valles, F. Design and Calibration of a Low-Cost SDI-12 Soil Moisture Sensor. *Sensors* **2019**, *19*, 491. [CrossRef]
6. Toure, S.; Diop, O.; Kpalma, K.; Maiga, A.S. Shoreline detection using optical remote sensing: A review. *ISPRS Int. J. Geo-Inf.* **2019**, *8*, 75. [CrossRef]
7. Scala, P.; Manno, G.; Ciraolo, G. Semantic segmentation of coastal aerial/satellite images using deep learning techniques: An application to coastline detection. *Comput. Geosci.* **2024**, *192*, 105704. [CrossRef]
8. Papakonstantinou, A.; Topouzelis, K.; Pavlogeorgatos, G. Coastline zones identification and 3D coastal mapping using UAV spatial data. *ISPRS Int. J. Geo-Inf.* **2016**, *5*, 75. [CrossRef]
9. Yuan, S.; Li, Y.; Bao, F.; Xu, H.; Yang, Y.; Yan, Q.; Zhong, S.; Yin, H.; Xu, J.; Huang, Z.; et al. Marine environmental monitoring with unmanned vehicle platforms: Present applications and future prospects. *Sci. Total Environ.* **2023**, *858*, 159741. [CrossRef]
10. Specht, M. Methodology for Performing Bathymetric and Photogrammetric Measurements Using UAV and USV Vehicles in the Coastal Zone. *Remote Sens.* **2024**, *16*, 3328. [CrossRef]
11. Mishra, M.; Chand, P.; Beja, S.K.; Santos, C.A.G.; da Silva, R.M.; Ahmed, I.; Kamal, A.H.M. Quantitative assessment of present and the future potential threat of coastal erosion along the Odisha coast using geospatial tools and statistical techniques. *Sci. Total Environ.* **2023**, *875*, 162488. [CrossRef]
12. Figliomeni, F.G.; Specht, M.; Parente, C.; Specht, C.; Stateczny, A. Modeling and Accuracy Assessment of Determining the Coastline Course Using Geodetic, Photogrammetric and Satellite Measurement Methods: Case Study in Gdynia Beach in Poland. *Electronics* **2024**, *13*, 412. [CrossRef]
13. Azad, S.; Neffati, A.A.; Mahmud, M.; Kaiser, M.S.; Ahmed, M.R.; Kamruzzaman, J. UDTN-RS: A New Underwater Delay Tolerant Network Routing Protocol for Coastal Patrol and Surveillance. *IEEE Access* **2023**, *11*, 142780–142793. [CrossRef]
14. Qiao, Y.; Yin, J.; Wang, W.; Duarte, F.; Yang, J.; Ratti, C. Survey of deep learning for autonomous surface vehicles in marine environments. *IEEE Trans. Intell. Transp. Syst.* **2023**, *24*, 3678–3701. [CrossRef]
15. Khan, A.R.; Ab Razak, M.S.B.; Yusuf, B.B.; Shafri, H.Z.B.M.; Mohamad, N.B. Future prediction of coastal recession using convolutional neural network. *Estuar. Coast. Shelf Sci.* **2024**, *299*, 108667. [CrossRef]
16. Halicki, A.; Specht, M.; Stateczny, A.; Specht, C.; Specht, O. Shoreline extraction based on LiDAR data obtained using an USV. *TransNav Int. J. Mar. Navig. Saf. Sea Transp.* **2023**, *17*, 445–453. [CrossRef]
17. Chen, S.; Huang, J.; Miao, H.; Cai, Y.; Wen, Y.; Xiao, C. Deep Visual Waterline Detection for Inland Marine Unmanned Surface Vehicles. *Appl. Sci.* **2023**, *13*, 3164. [CrossRef]
18. Zollini, S.; Dominici, D.; Alicandro, M.; Cuevas-González, M.; Angelats, E.; Ribas, F.; Simarro, G. New Methodology for Shoreline Extraction Using Optical and Radar (SAR) Satellite Imagery. *J. Mar. Sci. Eng.* **2023**, *11*, 627. [CrossRef]
19. Angelini, R.; Luzi, G.; Ribas Prats, F.; Masiero, A.; Mugnai, F. A review and test of shoreline extraction techniques. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2023**, *48*, 17–24. [CrossRef]
20. Park, J.; Kang, M.; Lee, Y.; Jung, J.; Choi, H.T.; Choi, J. Multiple autonomous surface vehicles for autonomous cooperative navigation tasks in a marine environment: Development and preliminary field tests. *IEEE Access* **2023**, *11*, 36203–36217. [CrossRef]
21. Luo, J.; Zhuang, J.; Jin, M.; Xu, F.; Su, Y. An energy-efficient path planning method for unmanned surface vehicle in a time-variant maritime environment. *Ocean Eng.* **2024**, *301*, 117544. [CrossRef]
22. Abd Alhattab, Y.; Abidin, Z.B.Z.; Faizabadi, A.R.; Zaki, H.F.; Ibrahim, A.I. Integration of Stereo Vision and moos-IVP for enhanced obstacle detection and navigation in unmanned surface vehicles. *IEEE Access* **2023**, *11*, 128932–128956. [CrossRef]
23. Petrovic, A.; Damaševičius, R.; Jovanovic, L.; Toskovic, A.; Simic, V.; Bacanin, N.; Zivkovic, M.; Spalević, P. Marine vessel classification and multivariate trajectories forecasting using metaheuristics-optimized extreme gradient boosting and recurrent neural networks. *Appl. Sci.* **2023**, *13*, 9181. [CrossRef]
24. Stateczny, A.; Delekta, M. Unmanned surface vehicles HydroDron-1 utilized in the MPSS project. *Innow. Metod. Badań Akwenów Portowych Z Wykorzystaniem Bezzałogowych Platf. Nawodnych* **2023**, 157–168. [CrossRef]

25. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. In Proceedings of the Advances in Neural Information Processing Systems 25: 26th Annual Conference on Neural Information Processing Systems 2012, Lake Tahoe, NV, USA, 3–6 December 2012. [CrossRef]
26. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778. [CrossRef]
27. Howard, A.G. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv* **2017**, arXiv:1704.04861.
28. Szegedy, C.; Vanhoucke, V.; Ioffe, S.; Shlens, J.; Wojna, Z. Rethinking the inception architecture for computer vision. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 2818–2826. [CrossRef]
29. Tan, M. Efficientnet: Rethinking model scaling for convolutional neural networks. *arXiv* **2019**, arXiv:1905.11946.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.