[Computers & Geosciences 192 (2024) 105704](https://doi.org/10.1016/j.cageo.2024.105704)

Contents lists available at [ScienceDirect](http://www.sciencedirect.com/science/journal/00983004)

Computers and Geosciences

journal homepage: [www.elsevier.com/locate/cageo](https://www.elsevier.com/locate/cageo)

[](http://crossmark.crossref.org/dialog/?doi=10.1016/j.cageo.2024.105704&domain=pdf)Semantic segmentation of coastal aerial/satellite images using deep learning techniques: An application to coastline detection

Pietro Scala , Giorgio Manno \*[, Giuseppe Ciraolo](#_bookmark1)

*Department of Engineering (DI), University of Palermo, Viale Delle Scienze, Building 8, 90128, Palermo, Italy*

A R T I C L E I N F O

*Keywords:*

Automatic coastline detection Deep learning

Coast train

Semantic segmentation Aerial/satellite images

A B S T R A C T

A new CNN based approach supported by semantic segmentation, was proposed. This approach is frequently used to carry out regional-scale studies. The core of our method revolves around a CNN model, based on the famous U-Net architecture. Its purpose is to identify different classes of pixels on satellite images and later to automatically detect the coastline. The recently launched Coast Train dataset was used to train the CNN model. Traditional coastline detection was improved (“water/land” segmentation) by means of two new aspects the use of the Sobel-edge loss function and the segmentation of the satellite images into several categories like built-up areas, vegetation and land besides beach/sand and water classes. The approach used ensures a more precise coastline extraction, distinguishing water pixels from all other categories. Our model adeptly identifies features, such as cliff vegetation or coastal roads, that some models might overlook. In this way, coastline localization and its drawing for regional scale study, have minor uncertainties. The performance of the CNN-based method, achieving 85% accuracy and 80% IoU (Intersection over Union) in the segmentation process. The ability of the model to extract the coastline was validated on a Sicilian case study, notably the San Leone beach (Agrigento). The model’s results align closely with the ground truth, moreover, its reliability was further confirmed when it was tested on other Sicilian coastal regions.

Beyond robustness, the model offers a promising avenue for enhanced coastal analysis potentially applicable to coastal planning and management.

# Introduction

Over the past few decades, there has been a notable increase in the frequency and severity of sea storms, which have had significant impacts on coastal areas, resulting in floods and erosion ([Castagno et al., 2018](#_bookmark40); [Emanuel, 2005](#_bookmark45), p. 201320132; [Lionello et al., 2017](#_bookmark58); [Murakami et al.,](#_bookmark66) [2017](#_bookmark66); [Stocker and Qin, 2013](#_bookmark76); [Vitousek et al., 2017](#_bookmark83); [Walsh et al., 2016](#_bookmark84)). The Intergovernmental Panel on Climate Change (IPCC) has highlighted these concerns, especially in its ([2022](#_bookmark51)) special report on the Ocean and Cryosphere. Coastal changes have led to extensive property loss, erosion, and damage to infrastructure ([Hinkel et al., 2014](#_bookmark50); [Satta et al.,](#_bookmark72) [2017](#_bookmark72); [Syvitski et al., 2009](#_bookmark78)).

Research has shown a rise in storm surge events, especially in the North Atlantic and Northwest Pacific regions, which will likely continue due to climate change ([Kopp et al., 2014](#_bookmark54); [Muis et al., 2020](#_bookmark65)). The Med- iterranean coast is also affected ([Lionello et al., 2017](#_bookmark58)), emphasizing the importance of understanding storm parameters and storm consequences when assessing coastal risks in the Mediterranean areas.

The proxy used to analyze the dynamics of the coast and to identify if the coast is in accretion or retreatment, due to storm surge events or sea level rise, is the coastline ([Pollard et al., 2020](#_bookmark69); [Sapkota and White,](#_bookmark71) [2019](#_bookmark71)).

Coastlines, the boundary between land and water, play a crucial role in studying coastal dynamics and are influenced by both natural events and human activities ([Anfuso et al., 2013](#_bookmark32), [2012](#_bookmark33); [Anfuso and Martínez](#_bookmark31) [Del Pozo, 2009](#_bookmark31); [Borzì et al., 2021](#_bookmark35)).

The coastline can be identified using various proxies like the top cliff line, seaward dune vegetation line, or wet/dry line, each with its pros and cons, depending on the study type. [Wu et al. (2019)](#_bookmark86) noted that coastlines are easily spotted where boundaries have sharp inclines, such as seawalls, or in areas with rough terrain like rocky or vegetated coasts. Additionally, the wet-dry line, accounting for wave and tide effects, can help in detecting coastlines ([Boak and Turner, 2005](#_bookmark34)).

Monitoring coastline changes is therefore essential for effective coastal management ([Manno et al., 2022b](#_bookmark63); [Molina et al., 2020](#_bookmark64)) but its perennial and considerable temporal variability makes it difficult to

\* Corresponding author.

*E-mail address:* [giorgio.manno@unipa.it](mailto:giorgio.manno@unipa.it) (G. Manno).

<https://doi.org/10.1016/j.cageo.2024.105704>

Received 28 August 2023; Received in revised form 8 May 2024; Accepted 15 August 2024

Available online 15 August 2024

0098-3004/© 2024 Elsevier Ltd. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

determine. In fact, the coastline position change at any instant, and this characteristic should be considered when satellite/aerial images are used ([Scardino et al., 2022](#_bookmark73)).

Satellite remote sensing provides continuous observations of coastal morphology, and the conventional method for detecting coastlines in- volves visual interpretation by experts. However, automated and semi- automated approaches are becoming more prevalent, and several methods have been developed using different instruments ([Toure et al.,](#_bookmark80) [2019](#_bookmark80); [Zhu et al., 2021](#_bookmark87))

Recent innovations have seen the rise of automated local-scale ap- proaches. These can be broadly categorized into boundary detection techniques ([Liu and Jezek, 2004](#_bookmark59); [Paravolidakis et al., 2018](#_bookmark68)) and seg- mentation methods ([Cao et al., 2020](#_bookmark39)). The latter often employs Machine Learning (ML) and Deep Learning (DL) strategies, such as Convolutional Neural Networks (CNNs) ([Aghdami-Nia et al., 2022](#_bookmark29); [Chang et al., 2022](#_bookmark41); [Kattenborn et al., 2021](#_bookmark53); [Sultan et al., 2020](#_bookmark77)). For instance, studies by [Dang et al. (2022)](#_bookmark43) and [Chang et al. (2022)](#_bookmark41) successfully used deep learning models to detect coastlines using high-resolution satellite im- agery. [Seale et al. (2022)](#_bookmark74) also employed segmentation methods to distinguish between sea and land.

However, detecting coastlines from satellite imagery is challenging due to various environmental characteristics and the problems of “poor spatiotemporal generalization” and “scaling” ([Liu et al., 2017](#_bookmark60); [Toure](#_bookmark80) [et al., 2019](#_bookmark80)). Addressing these issues, the current research uses a CNN model grounded in semantic segmentation for coastline identification, leveraging the segmentation technique differently from existing methods. This approach discerns the land-sea boundary and also detects human-made structures, vegetation, and other land cover. The research uses the Sobel-edge loss function from [Seale et al. (2022)](#_bookmark74) and employs the Coast Train dataset for model training ([Buscombe et al., 2023](#_bookmark37); [Wernette et al., 2022](#_bookmark85)). In fact, we trained the model with a dataset specific to coastal areas thus building a specialized classifier for these types of areas.

The manuscript is structured in five sections: introduction, materials and methods, Sicilian case studies, application of the method at the Sicilian coast, and conclusions.

# Materials and method

This section provides an overview of the data utilized and the methodology employed for detecting and delineating the coastline. To identify the coastline from satellite/aerial images, an automatic recog- nizing approach was used. In detail, a neural network was used, and its training and validation was made by means of the Coast Train dataset. The post-processing of the model output for the extraction of the coastline is also described. Finally, a description of the study area and the images used to test the model on the Sicilian case study is provided. The flowchart in [Fig. 1](#_bookmark3) describes the working principle of the pro-

posed model.

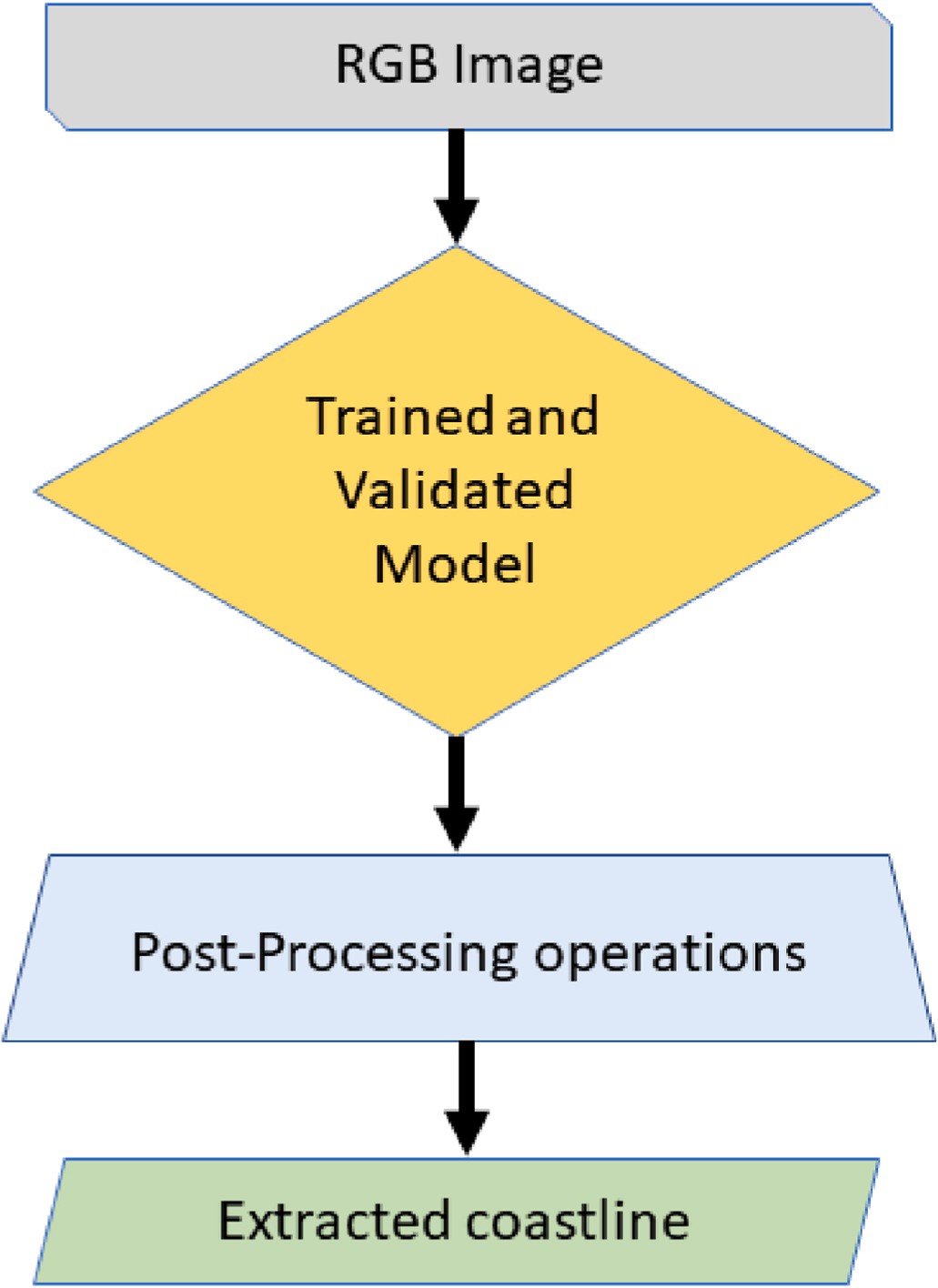
* 1. *The “Coast Train” dataset*

The Coast Train dataset ([Wernette et al., 2022](#_bookmark85)) includes 10 coastal images with 1.2 billion labeled pixels covering 3.6 million hectares. These free-to-use images vary in resolution, with aerial orthophotos

ranging from 0.05 m to 1 m and satellite images from 10 m to 15 m. The dataset, which covers regions from 26 to 48◦ North latitude and 69–123◦ West longitude, includes diverse coastlines from the Pacific, Atlantic,

and Gulf coasts. Features include rocky cliffs, wetlands, and sandy beaches in both rural and urban settings across different energy levels ([Buscombe et al., 2023](#_bookmark37)).

Sources for the dataset include NAIP aerial imagery (2010–2018), Sentinel-2 (2017–2020), Landsat-8 (2014–2020), USGS aerial imagery (2008–2012), and UAS orthomosaic imagery (2008–2012). Coastal ex- perts classified the images into broad categories, with all datasets labeled at least for "water" and "sand". For accuracy, only high-



**Fig. 1.** Model operation flowchart.

confidence data was annotated. The labeled data, organized into four to twelve categories, reflects the image features ([Buscombe et al., 2023](#_bookmark37)). The Coast Train dataset provides an invaluable resource to study coastal environments, such as deltas, estuaries, barrier islands, salt marshes, etc. This data can be downloaded from [https://www.usgs.gov/data/coast-tr](https://www.usgs.gov/data/coast-train-labeled-imagery-training-and-evaluation-data-driven-models-image-segmentation) [ain-labeled-imagery-training-and-evaluation-data-driven-models-imag](https://www.usgs.gov/data/coast-train-labeled-imagery-training-and-evaluation-data-driven-models-image-segmentation) [e-segmentation](https://www.usgs.gov/data/coast-train-labeled-imagery-training-and-evaluation-data-driven-models-image-segmentation) in which all data specifications are also described. To apply, we carefully selected a dataset of 134 JPEG images and their corresponding label masks. To ensure the highest possible quality, we began by removing any images that did not clearly depict a land-water interface. We then specifically selected images that closely resembled the coastal environments of Sicily, Italy, and limited our selection to a maximum of 10 images per environment, each with varying sea states. By employing this rigorous selection process, we were able to construct a final dataset that is both relevant and diverse, providing an excellent base for our research.

* 1. *The “Coast Train” dataset augmentation*

To avoid the issue of overfitting the model and to obtain a larger dataset, the authors utilized data augmentation on the Coast Train im- ages and their masks. The CNN model reads the images as a three- dimensional array of numbers, representing the color values of each pixel. By using the data augmentation technique, these pixels can be modified to generate new, augmented images that resemble the original dataset, but with additional information to improve the generalization of the ML algorithm.

Overfitting arises when a model excels on the training data but underperforms on previously unencountered data, and this can happen when the dataset is small. To prevent this from happening, the authors

used the Python library *Albumentations* ([Buslaev et al., 2020](#_bookmark38)) for data augmentation, that employs an array of image transformation tech- niques fine-tuned for optimal performance.

The original dataset consisted of 134 images, with 100 used for training and 34 for validation. However, given the limited size of the dataset, data augmentation techniques were employed to synthetically expand the amount of training data by a factor of four. The final training dataset consisted of 502 images and the validation dataset remained at 34 images. The data augmentation processes used were random crop- ping, horizontal and/or vertical flipping, rotation, adjustments to brightness and contrast, and grid distortion.

* 1. *The U-Net model and used loss function*

U-Net is a renowned convolutional neural network (CNN) initially designed for biomedical image segmentation tasks ([Ronneberger et al.,](#_bookmark70) [2015](#_bookmark70)). In the study, it was used for sand/sediment, built-up, vegetation, land and water/sea segmentation at the pixel level. Each pixel without segmentation is labeled as “No Label”.

The U-Net framework is structured around two main components: the encoder and the decoder.

The encoder can be likened to a contracting pathway and bears resemblance to a conventional CNN. Each segment within the encoder consists of a pair of 3 x 3 convolutional layers, which are subsequently followed by a Rectified Linear Unit (ReLU) activation function. To wrap up each block, a max-pooling layer with a 2 x 2 kernel and a stride of 2 is incorporated. The inherent nature of the encoder is to methodically reduce tensor dimensions while amplifying feature channels, aiming to capture intricate low-level image details. The consequence of this is that the number of feature channels doubles after every encoder block, preserving the rich content of the image within a compact feature vector.

The decoder work as an expansive pathway. Each segment of the decoder comprises up-sampling procedures, concatenation layers, and two 3 x 3 convolutional layers. Its task is to methodically expand the tensor dimensions back to their original metrics, while superimposing the valuable features extracted by the encoder onto their corresponding spatial regions. Throughout each decoder block, the tensor dimensions double courtesy of the up-sampling mechanism.

The concluding phase of the architecture utilizes a 1 x 1 convolution applied to the decoder’s output. This serves to map the feature vectors of individual pixels to the requisite number of classifications, culminating in a pixel-by-pixel segmentation mask. Notably, the values within this map quantify the likelihood of a given pixel’s affiliation to a particular class.

[Fig. 2](#_bookmark5) illustrates the U-Net model’s architecture.

Choosing an appropriate loss function for coastline segmentation is

crucial. It directs the training process to determine the best parameters and plays a role in how these parameters are adjusted ([Galeone, 2019](#_bookmark47)).

For this reason, Sobel-edge-loss was used as the loss function ([Seale](#_bookmark74) [et al., 2022](#_bookmark74)). A more detailed description presented in the paper by [Seale](#_bookmark74) [et al. (2022)](#_bookmark74) and used in this paper can be found in [Appendix 1](#_bookmark24).

* 1. *Architecture settings*

Within the framework of our proposed architecture, we have incor- porated various hyperparameters that influence both the training pro- cess and the overall performance of the model. These parameters, which are external to the model itself, are described below.

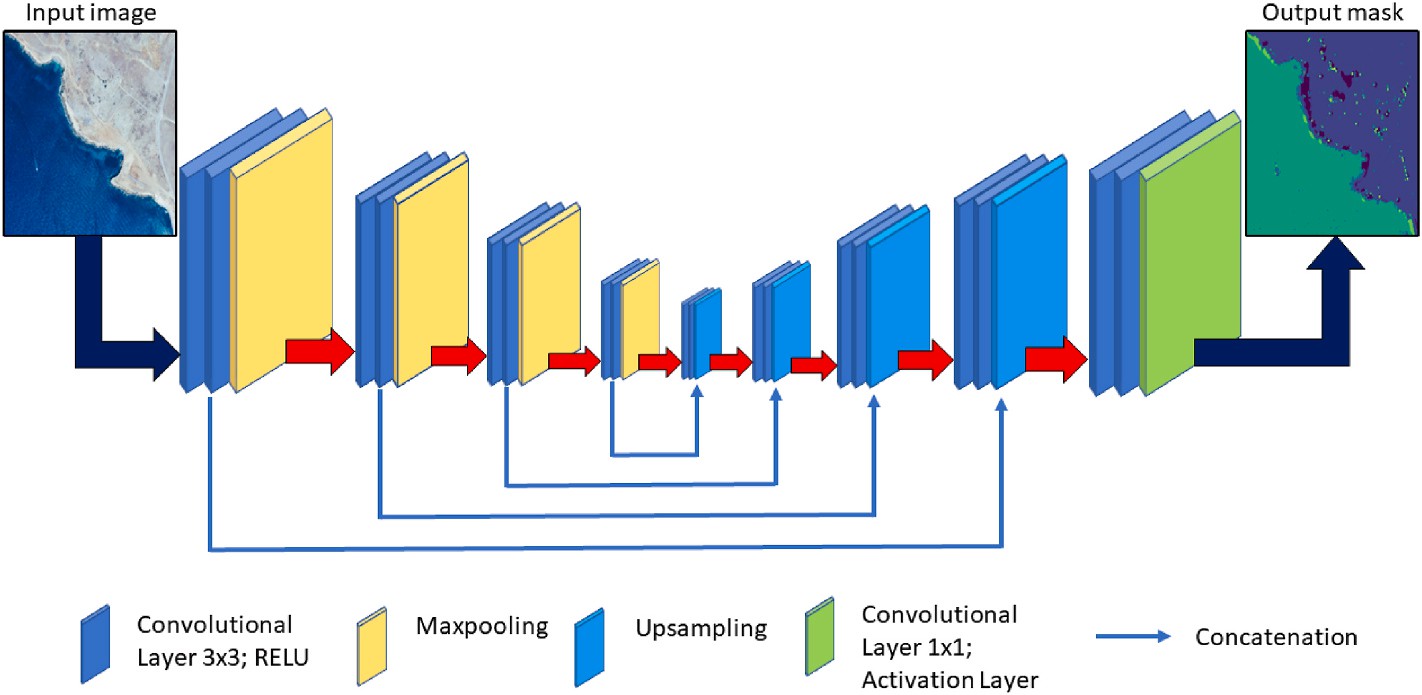
The batch size is the number of data samples used to compute a single iteration during the training process. A larger batch size may lead to faster training but could require more memory. In our model, the batch size was set to 16 according to [Kandel and Castelli (2020)](#_bookmark52). The number of epochs represents the number of complete passes through the entire dataset during the training process. While a higher number of epochs may promote better learning of the training data, there is a risk of overfitting if not controlled. In our case, the maximum number of epochs was set to 60. Regarding the optimizer, the Adam algorithm was employed, which relies on adaptive estimates of first and second-order moments to iteratively update the model weights to minimize the loss function. During the training process, three different callbacks were utilized:

1. EarlyStopping: this callback stops training if the loss on the evalua- tion set does not improve for a specific number of consecutive epochs (patience), which in our case was set to 10 epochs.
2. ReduceLROnPlateau: this callback reduces the learning rate if no improvement is observed in the monitored evaluation metric. The chosen metric is evaluation accuracy. In addition, a learning rate reduction factor of 50% (0.5) was defined each time the metric does not improve.
3. ModelCheckpoint: this callback saves the model with the lowest loss during training, ensuring the preservation of the best-performing model.

A lower bound on the learning rate was set at 10-5 to prevent further reductions when the learning rate reaches or falls below this value

([Kandel and Castelli, 2020](#_bookmark52)). This maintains a sufficiently small learning rate to facilitate convergence or prevent oscillations around the global minimum of the loss function.

We also utilized an Input Size of 512 x 512 pixels. According to recent studies ([Lin et al., 2019](#_bookmark57); [Nekrasov et al., 2018](#_bookmark67); [Tanveer et al.,](#_bookmark79) [2022](#_bookmark79)), this resolution enhances model training and accuracy, enabling better differentiation between land, sea, and other classes. A 512 x 512



**Fig. 2.** Diagram framework illustration of the U-Net architecture.

resolution, compared to 256 x 256, provides more detail and context. It captures fine details and allows the model to consider spatial relation- ships between objects, improving accuracy ([Lin et al., 2017](#_bookmark56)). This res- olution also reduces aliasing, where high-frequency information is misrepresented ([Zou et al., 2023](#_bookmark88)), and the border effect, critical for distinguishing between water and land, is enhanced at this resolution ([Dhingra et al., 2021](#_bookmark44)). Moreover, despite the higher computational de- mand, models trained at this resolution show greater robustness and generalize better to new data ([Nekrasov et al., 2018](#_bookmark67); [Shi et al., 2021](#_bookmark75)).

To confirm all these aspects, we conducted a sensitivity analysis on the performance of the model by training it with a smaller input size, 256 x 256. Lower sizes were not taken into account considering the nature of coastal images that would be poorly represented by resolutions smaller than 256 x 256. The results confirmed that the optimal input size is 512 x 512 and are shown in [Appendix 3](#_bookmark28).

Regarding class weights, during model training, different weights are usually assigned to classes based on their frequency in the dataset. This procedure ensures that the model gives more importance to the less represented classes during the optimization of the loss function ([Fer-](#_bookmark46) [nando and Tsokos, 2021](#_bookmark46); [Tyagi and Mittal, 2020](#_bookmark81)). However, we did not adopt this practice, but considered that each class contributes equally to the overall loss of the model during the training process. This was because the initial training dataset was already balanced and the data augmentation was performed during preprocessing to avoid computa- tional overhead in training the classification model. Since the main goal was to achieve accurate segmentation of water boundaries with respect to other classes, we accepted this compromise.

The model was developed using TensorFlow 2.11 and Python 3.9. All training tasks were executed on a Desktop Computer powered by Win- dows 11 Pro, an Intel® Core™ i7-12700 12th Gen processor, 16 GB DDR4 3200 MHz RAM, Intel 770 UHD onboard graphics, and a 1 TB nvme SSD storage. In our work, the semantic segmentation model was executed on a dedicated GPU to ensure optimal performance and computational efficiency.

* 1. *Model training and validation*

To assess how well the model classifies, we employed the Accuracy ([Albon, 2018](#_bookmark30); [Brownlee et al., 2021](#_bookmark36)) and Intersection over Union (IoU) or Jaccard index ([Girshick et al., 2014](#_bookmark48); [He et al., 2018](#_bookmark49); [Ronneberger](#_bookmark70) [et al., 2015](#_bookmark70)), as metrics.

In computer vision and image processing fields, two key metrics used to gauge the effectiveness of object detection and image segmentation algorithms are Accuracy and Intersection over Union (IoU), commonly referred Jaccard index.

Accuracy is a straightforward metric, primarily gauging the ratio of accurate predictions to the overall number of predictions. This value lies between 0 and 1, with 1 indicating impeccable accuracy and 0 denoting absolute inaccuracy.

Conversely, IoU is a nuanced metric aimed at assessing the degree of overlap between the forecasted regions and the actual bounding boxes or segmentations. This metric also varies between 0 and 1, where 1 signifies an exact match and 0 indicates no intersection. IoU is especially pertinent for tasks that emphasize pinpointing the precise region of an object within an image. The mathematical expressions for these metrics are described in appendix section 2.

We have also trained a widely used architecture in the field of se- mantic segmentation, DeepLabV3 ([Chen et al., 2017](#_bookmark42)) with the same training and validation dataset used in UNet to compare the classifica-

tion ability of the architecture proposed in this work against DeepLabV3.

DeepLabv3 ([https://github.com/AvivSham/DeepLabv3?tab=readm](https://github.com/AvivSham/DeepLabv3?tab=readme-ov-file) [e-ov-file](https://github.com/AvivSham/DeepLabv3?tab=readme-ov-file)) unlike UNet is a deep neural network (DNN) architecture widely used in semantic segmentation. It uses Atrous (Dilated) Convo-

lutions to control the resolutions of the receptive field and feature map without increasing the total number of parameters. In addition, through the use of Atrous Spatial Pyramid Poolingl, the architecture is able to

effectively extract multiscale features that contain useful information for segmentation. In general, the network is able to capture dense feature maps with rich long-range information that can be used to accurately segment images. Because of these features, the DeepLab V3 architecture is an excellent performance comparison tool for the model presented in this paper.

[Fig. 3](#_bookmark6) shows a diagram illustrating the entire procedure of creating data and instructing the neural network model. By giving as input an image depicting a coastal environment, the pre-trained model provides as output a mask image (prediction image) of the same environment.

* 1. *Post-processing*

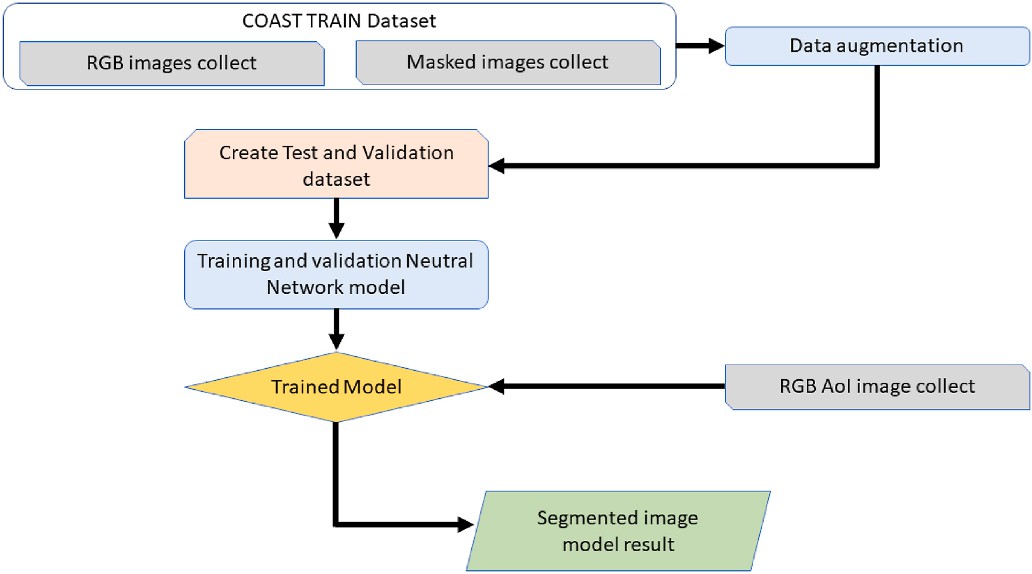
Once the coastal prediction image is obtained from the model, it is subjected to an image thresholding process. In this work we used the global thresholding: in fact a unique threshold value is used for the entire image, which can result in loss of information because the image has multiple regions with different intensity values.

The image is thus filtered with respect to all pixels that have RGB band values lower than those corresponding to the mask of the ’sedi- ment’ category representing the beach or cliff, thus leaving out the plume regions of the sediment. The intermediate pixels result as a kind of disturbance in the image. A conversion to unsigned bytes is then applied to the thresholding result, followed by a mean blurring opera- tion to eliminate the disturbance of the intermediate pixels and maintain the edges.

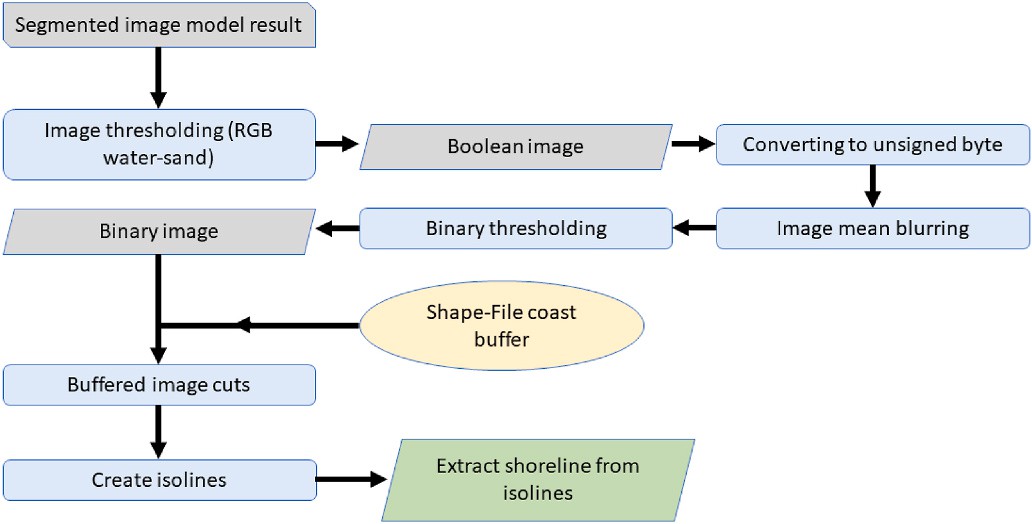
Mean blurring is a technique that smoothens images by reducing noise or intricate details. This method calculates the average value of nearby pixels and uses it to replace the original pixel value. This is achieved by using a kernel, a small square-shaped matrix, where the numbers dictate the significance of adjacent pixels. The extent of the blur is determined by the kernel’s size. A larger kernel means a broader average, leading to more pronounced blurring. In our study, we utilized a kernel with a size of 5. After this, binary thresholding is employed to divide the image into two clear categories. The result of this binary thresholding is a binary image, or a black and white image, which is used for the coastline detection. This image is then cropped to reduce its size, and thus computation time, by applying a 50-m buffer of a probable coastline indicative in shape-file vector format (SHP) of the area under consideration.

The coastline is not determined by the edge between black and white pixels but by an isoline within the pixels that passes through the threshold value. Buffer clipping and shoreline extraction operations are performed by exploiting QGis software. This method, proposed by Doan (2021), creates a coastline that resembles a natural coastline with a curved shape instead of a rough boundary between pixels.

The above-mentioned post-processing method is performed following the flowchart in [Fig. 4](#_bookmark7).



**Fig. 3.** Flowchart of the procedure to create data for train and validate the convolutional neural network model to identify segmented images.



**Fig. 4.** Flowchart of the post-processing process for identifying coastlines using convolutional neural network model image result.

# Study area and related dataset to apply the model

There are several morphotypes along the coasts of Sicily. The north- eastern Tyrrhenian sector and northern Ionian coasts are made up of headlands and pocket pebble beaches. The southern Ionian coast the landscape changes into sandy beaches, low rocky coast interrupted by cliffs. Finally, long sandy beaches and cliffs coastal sectors characterize the Mediterranean coast. The influence of human activity on the coastline is evident along the entire coastline. As stated in [Manno et al.](#_bookmark61) [(2022a)](#_bookmark61), the Sicilian coastline was divided into 22 coastal cells referred to as "Littoral Cells (LC)" ([Fig. 5](#_bookmark9)). Littoral cell is defined as that portion of the coastal territory that has no exchange of sediments with neighboring cells. In this study, we used the term "Littoral Cells" specifically for the second-degree LCs. These LCs have both natural and anthropic divisions, such as large ports or marine structures. Each LC is designated by a pair of numbers: the initial number indicates the first-order LC, and the subsequent one represents the second-order LC ([Manno et al., 2022a](#_bookmark61)).

For the coastline extraction accuracy analysis, we applied the model on the coastline of San Leone in Sicily (Italy).

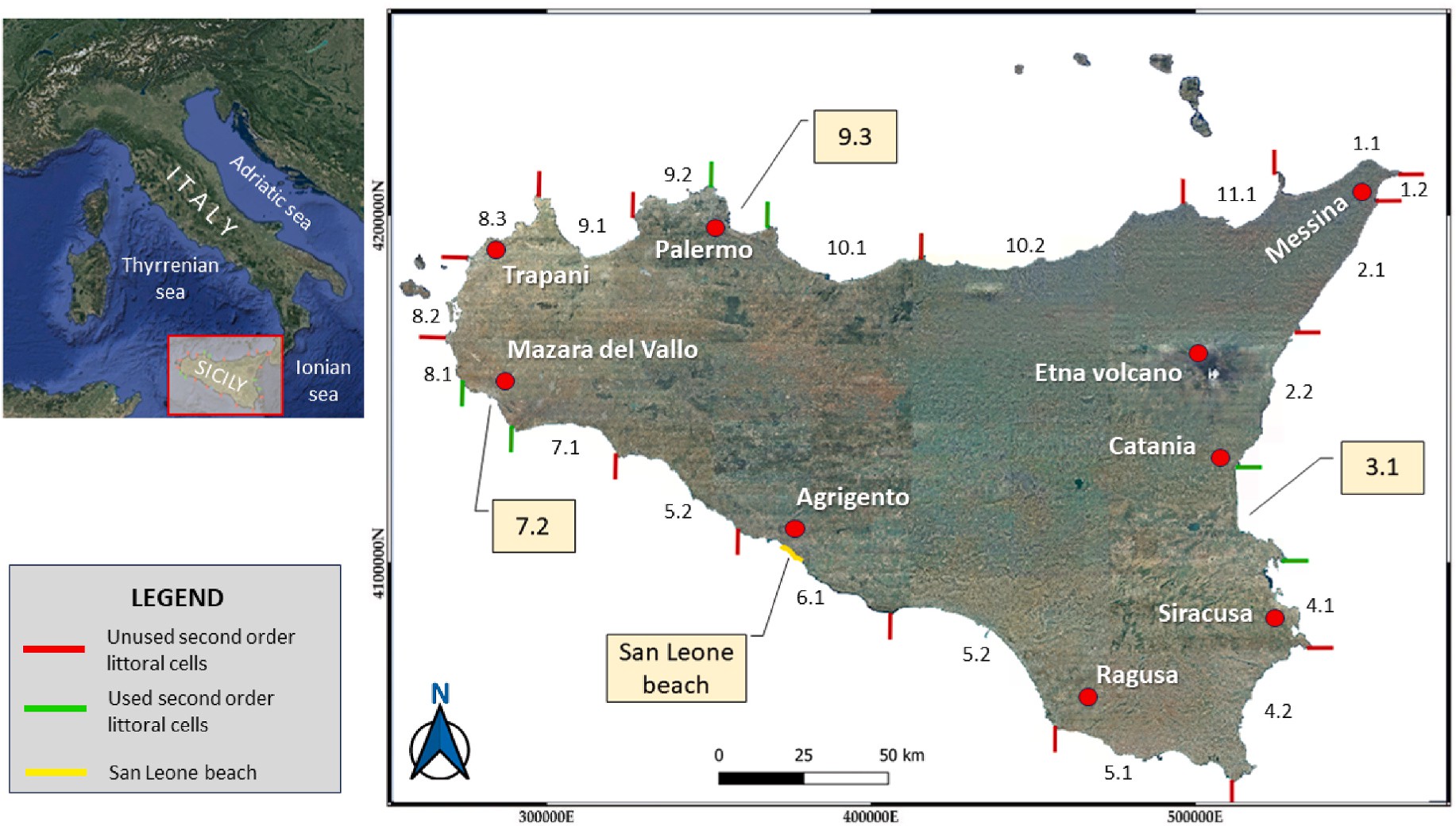
The beach of San Leone (near to Agrigento) is a part of the Sicilian Secondary LC no. 6.1 (Capo Bianco-Lido Rossello) and stretches for about 9 km including the mouths of the river Naro. The beach at San

Leone is composed of golden sand and has sections with dunes of varying heights (3–4 m). Over the last decades, the coast of San Leone has un- dergone uncontrolled development, turning the fishing village into a chaotic seaside resort. The beach is divided into three sub-areas, each with specific characteristics. Along the coast is located also the tourist- fishing harbour of San Leone, which has become the most important coastal structure in the area. The presence of structures such as break- waters and gabions along this stretch of coastline has had an impact on the morphology of the beach and the coastline, causing significant changes ([Manno and Ciraolo, 2015](#_bookmark62)). This beach was chosen because it contains several elements with which to efficiently test the detection model, after the training and validation phase. These elements are represented by the presence of a harbour, coastal protection structures, the presence of an estuary as well as the possibility of testing the behaviour of the model on a sufficiently extended beach. For this reason, the coastline of the study area, besides encompassing the previously mentioned San Leone coastline, extends for a further 1.8 km south of the mouth of the River Naro with a beach with similar characteristics to those of San Leone and without the presence of anthropic structures. This coastline was obtained by digitizing the 1: 10,000 map sheets of the ATA 2000 flight and can be consulted at [http://www.pcn.minambiente.](http://www.pcn.minambiente.it/mattm/progetto-coste/) [it/mattm/progetto-coste/](http://www.pcn.minambiente.it/mattm/progetto-coste/).

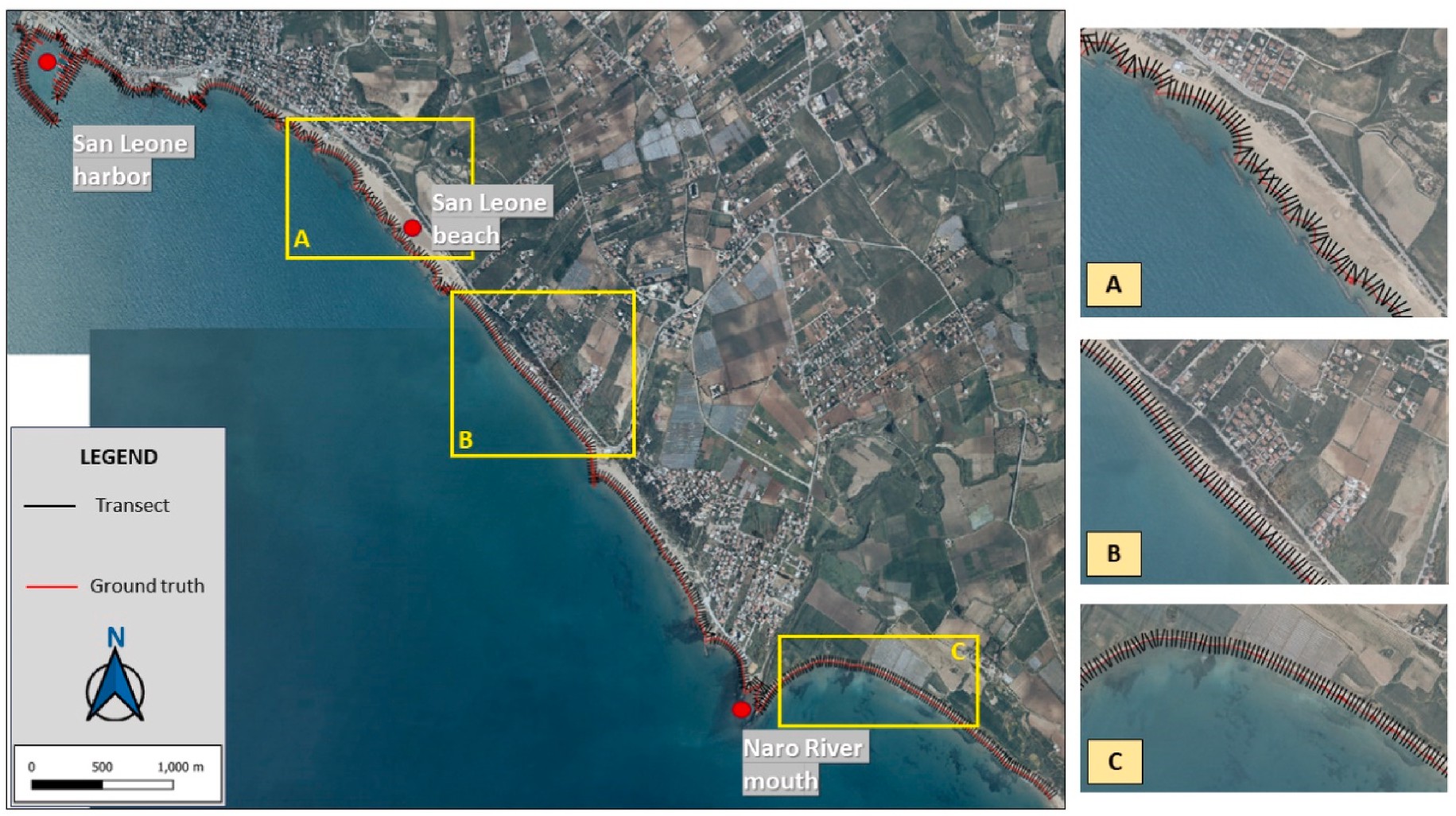
For this further goodness-of-fit analysis of the model, transects were

created orthogonal to the ground truth line with an average inter- distance of 6 m and 30 m in length ([Fig. 6](#_bookmark10)). The transects have a smaller spacing in stretches where the ground truth line appears to have little variation in shape while the presence of transects appears to be denser where the ground truth coastline has particular undulations or variations in shape. The total number of transects analyzed is 1887.

Once we further tested the model with images representing a het- erogeneous stretch of the Sicilian coastline, we extracted coastlines from a complete and diverse range of coastal environments (such as the presence of harbors, cliffs and beaches, anthropic defence structures, etc.). In detail the LCs 3.1 (Ionian coast), 9.3 (Tyrrhenian coast), and 7.2 (Mediterranean coast) were chosen. In [Fig. 5](#_bookmark9), the Sicilian macro-coastal sectors divided into second-level LCs and the chosen cells can be seen. For the detection of these coastlines orthophotos from the Italian *Geo- portale Nazionale*, taken during the 2013 ATA flights, in the three chosen



**Fig. 5.** Breakdown of Sicilian macro-coastal regions into second-level LCs. The 2013 ATA flight images of coastlines between the green segments are extracted in this work. (Reference System: WGS84-UTM33N-EPSG: 32633).



**Fig. 6.** Transects (segments in black) perpendicular to the ground truth line (red line) in the first column of the plot. The transects have an average inter-distance of 6 m and a length of 30 m. Sub-plot A, B and C show enlargement of the San Leone area (second column).

coastal regions of Sicily were utilized. Each LC is divided in 17, 8 and 11 1:10,000 scale images for 3.1, 9.3 and 7.2 LC respectively, with the 512 x 512 number of pixel image resolution.

# Results and discussion

This section delves into the analysis and discussion of outcomes achieved during the model’s training and validation phases, based on the U-Net architecture. The focus is on the accuracy, IoU, and loss metrics as detailed in section [2.4](#_bookmark4). In this section, the aforementioned metrics are analyzed and discussed on a Sicilian case study and the detected case study coastline is compared with a validated coastline representing the ground truth.

The efficiency of the proposed coastline detection approach in three different Sicilian coastal areas (LCs of section [3](#_bookmark8)) is also examined, and the detection performing for the different littoral zones is analyzed.

* 1. *Model metrics*

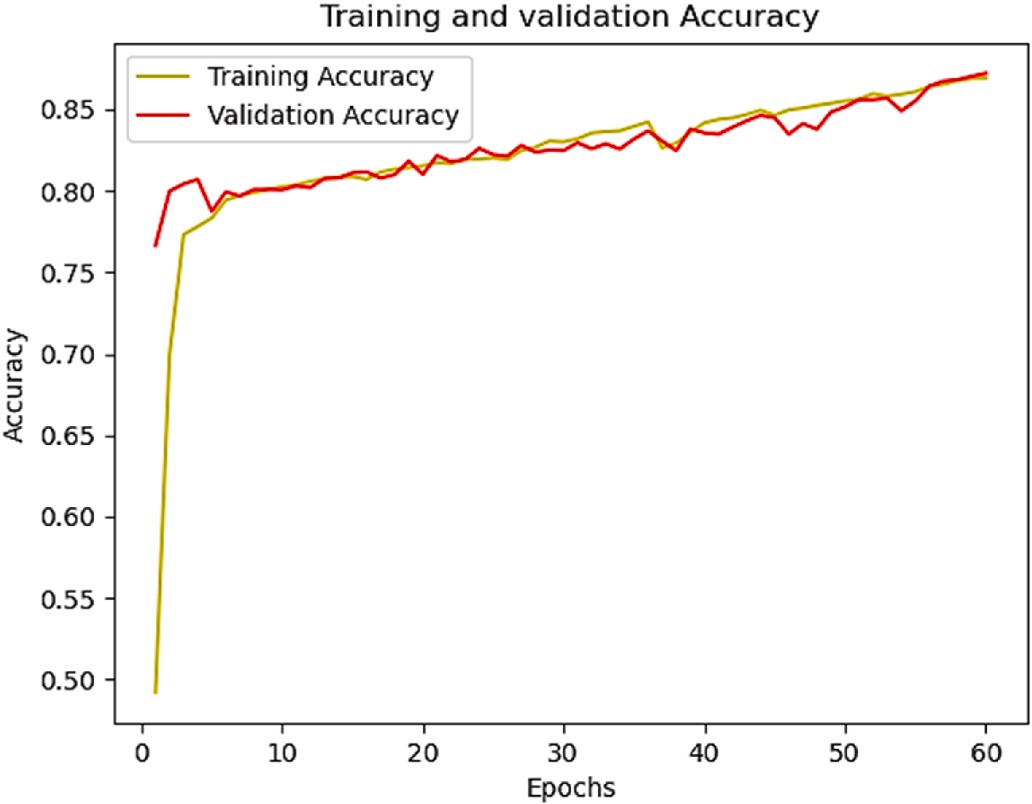
The model performance was automatically evaluated by the metrics described in Section [2.4](#_bookmark4). Our model based on the U-Net structure ach- ieved an accuracy of more than 80 percent and reaching the maximum value around 87 percent.

Analysis of the training and validation accuracy plot ([Fig. 7](#_bookmark12)) high- lights some interesting features of the semantic segmentation model during the training process.

First of all, it can be noted that the accuracy curves during training and validation processes show steady and gradual growth in the early epochs, indicating that the model is effectively learning the patterns of the different classes in the image.

However, from the fifth epoch onward, the curves stabilize around 80–87 percent accuracy with less steep growth for both training and validation. This suggests that the model has reached a saturation point and may have difficulty improving performance further. It is possible that the complexity of the problem or the image dataset used limit the potential for improvement beyond this point.

Another important observation is that the validation accuracy curve remains slightly below the training accuracy curve throughout the



**Fig. 7.** Training and validation model Accuracy plot.

training process. This discrepancy could suggest the presence of slight overfitting, i.e., the model has learned to classify training data correctly, but may not generalize as well to new data not seen in training.

In general, the trend of the curves in the graph indicates that the model achieved a good accuracy around 80 percent already in the early epochs and maintained a steady but modest growth in the later epochs. However, it is also important to evaluate other metrics such as Loss and IoU to get a complete view of the performance of the semantic seg- mentation model and avert overfitting issues.

[Fig. 8](#_bookmark13) shows the trend of IoU, which appears to be similar to that of accuracy. This indicates that the two measures are correlated, and that the semantic segmentation model is achieving good results in segmen- tation accuracy.

A high IoU indicates that the segmentation predicted by the model overlaps significantly and accurately with the reference segmentation. If the graph of IoU follows a similar trend to that of accuracy, it therefore



**Fig. 8.** Training and validation Jaccard coefficient plot (IoU).

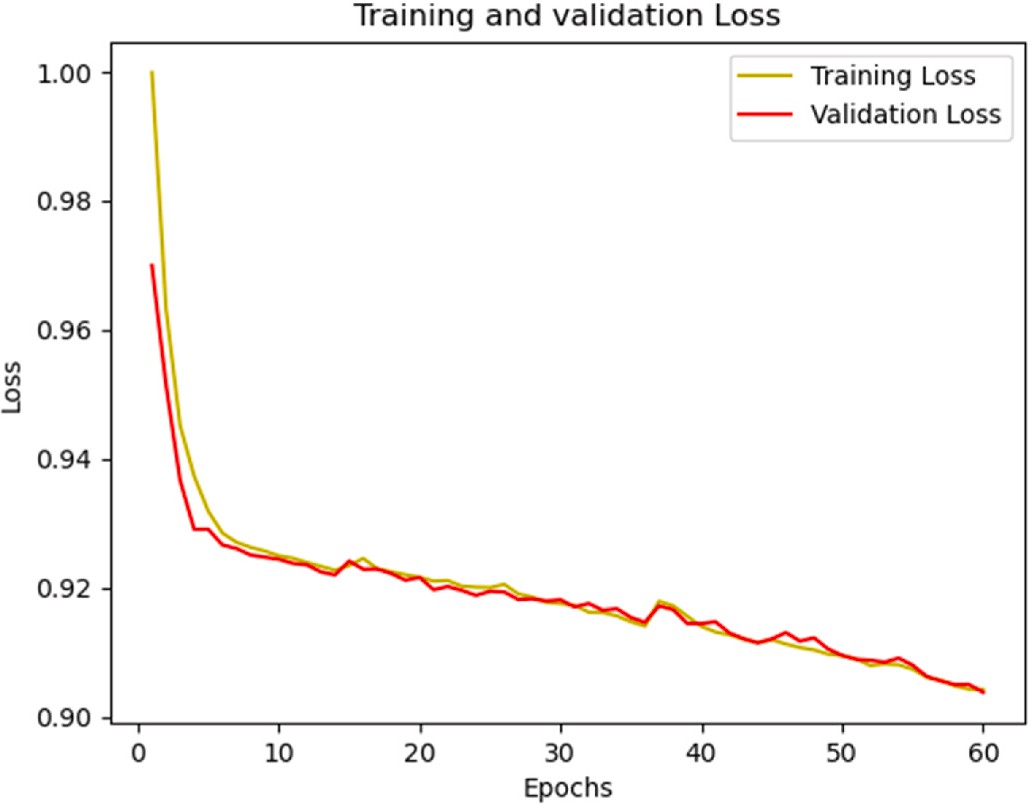
means that the increase in accuracy is accompanied by an improvement in IoU.

In general, the IoU graph shows high growth in the first 5 epochs, reaching a maximum value of 80%. However, the growth becomes slower in the later epochs. Unlike the accuracy graph, here the valida- tion curve is not always below that of training. This indicates that the model is generalizing well to new data during the validation phase, maintaining similar performance to that of training. The maximum IoU achieved of 74 percent suggests that the model is able to generate ac- curate segmentations that overlap significantly with the reference segmentations.

This is therefore a positive fact, since IoU is a more specific measure for assessing the quality of semantic segmentation than accuracy, which may be affected by unbalanced classes or dominant background regions. However, it is important to note that although IoU and accuracy show a similar trend, IoU provides more detailed information about the quality of segmentation, as it also takes into account the overlap of

specific classes.

Finally, in [Fig. 9](#_bookmark14) we show the graph of loss which shows a strong decrease in the first 5 epochs, from 1.015 to 0.93, and then continues to decrease slowly until the 60th epoch to the minimum value of 0.904. As the training and validation curves almost overlap, the figure indicates



**Fig. 9.** Training and validation model Loss plot.

that the model is well generalized and maintains good predictive ability even on new data. The decrease in Loss indicates that the model is gradually learning to generate more accurate and consistent predictions during the training process on new examples.

Overall, graph analysis suggests that the model achieves good ac- curacy and IoU in semantic segmentation relatively early in the training process. However, the subsequent growth is slower, suggesting that the model’s performance has reached a saturation point. The validation curve is generally slightly lower than the training curve, but the dif- ference being small, thus indicating good generalization ability.

We also compared the training and validation metrics with those obtained from the DeepLabV3 network trained with the same pre- processed Coast Train dataset (see Section 1.5). The results are shown in [Table 1](#_bookmark15).

From [Table 1](#_bookmark15), shows that although the DeepLAbV3 metrics are found to be acceptable for classifiers of this type, however, the U-Net archi- tecture provides slightly better performance especially in the validation phase. The results obtained suggest that the UNet architecture is able to generate (albeit slightly) more accurate predictions and has a greater ability to correctly segment coastal images than DeepLabv3. In fact, The U-Net model is known to be efficient in capturing spatial details due to its U-shaped structure, which allows it to preserve detailed information during the encoding and decoding process. On the other hand, Deep- Labv3 which uses Atrous convolution to expand the field of view and capture contextual information at different scales, might be less suitable in purely coastal image segmentation with few segmentation classes, falling into overfitting. Results on validation suggest this behavior. Similar results were obtained and reported in Gonzalz-Perez et al. (2022).

* 1. *Model’s performances on the Sicilian case study: San Leone beach*

Since the model’s performance evaluation was conducted exclu- sively on the dataset imported for training and validation (Coast Train), it’s unfeasible to gauge the model’s effectiveness on one dataset by using the metrics of a different dataset. Moreover, since the coastline is derived from isolines, evaluating accuracy purely based on pixels in the pre-processing phase would be both imprecise and potentially misleading. To further evaluate the quality of the model after training, the model was used to define the coastline on orthophotos different from those in the training and validation dataset.

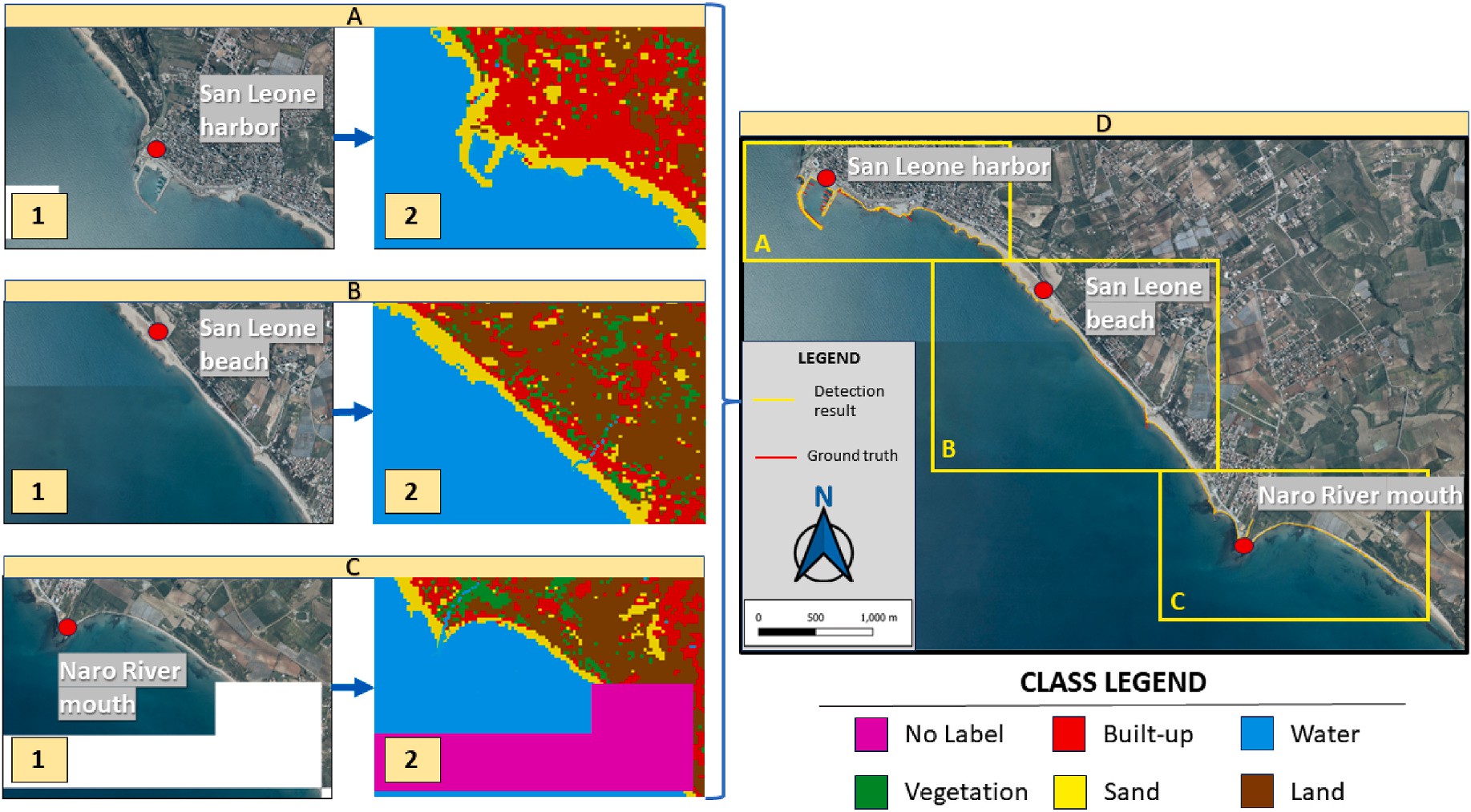
The images A1, B1, C1 in [Fig. 10](#_bookmark16) were used as input to the model (flowchart in [Fig. 1](#_bookmark3)), which provided graded images (A2, B2, C2 in [Fig. 10](#_bookmark16)) as output. Note how the pixels in the image belonging to the port facilities are classified as sand. The color of these pixels is in fact very similar to that of the beach. However, since the interest is to make the detection of the coastline, the classification is deemed acceptable. After post-processing operations, the coastline was extracted (panel D in [Fig. 10](#_bookmark16)). The entire coastline extraction process averages less than 1 min of processing time.

The coastline extracted from the model and the one for comparison, representing ground truth are shown in yellow and red lines respectively in [Fig. 10](#_bookmark16) (panel D). This coastline represents the ground truth of comparison and denotes the transition line between water and land/ sand. [Fig. 10](#_bookmark16), panel D, shows for the entire coastline section analyzed,

**Table 1**

Comparison of metrics (Accuracy and IoU) for the proposed architecture and DeepLabV3, both trained on the same dataset. In bold the best values of the metrics.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Architecture | Training Accuracy | Validation Accuracy | Training IoU | Validation IoU |
| *Proposed CNN* | **0.87** | **0.87** | 0.74 | **0.75** |
| *architecture*  *DeepLabV3* | 0.85 | 0.83 | **0.76** | 0.73 |



**Fig. 10.** San Leone coastline detection. Panels A1, B1, C1 represent the merge of input images provided to the model. Panels A2, B2 and C2 represent the merged output of the model (classified images). Panel D shows the overlay between the ground truth coastline (red line) and the coastline extracted from post processing operations (yellow line). The yellow boxes of panel D represent orthophotos A1, B1 and C1.

about 11 km, that the two lines tend to blur suggesting good detection by the model.

Each transect built on the ground truth shoreline in [Fig. 6](#_bookmark10) (see section [3](#_bookmark8)) was cut at the two coastlines in [Fig. 10](#_bookmark16), and its length was calculated for each resulting section. Cut transects with lengths greater than 1 m are shown in red and transects with lengths less than 1 m are shown in green in [Fig. 11](#_bookmark17).

[Fig. 12](#_bookmark18), panel A, shows a histogram representing the number of cut transects having a length less than 1 m and the number of cut transects

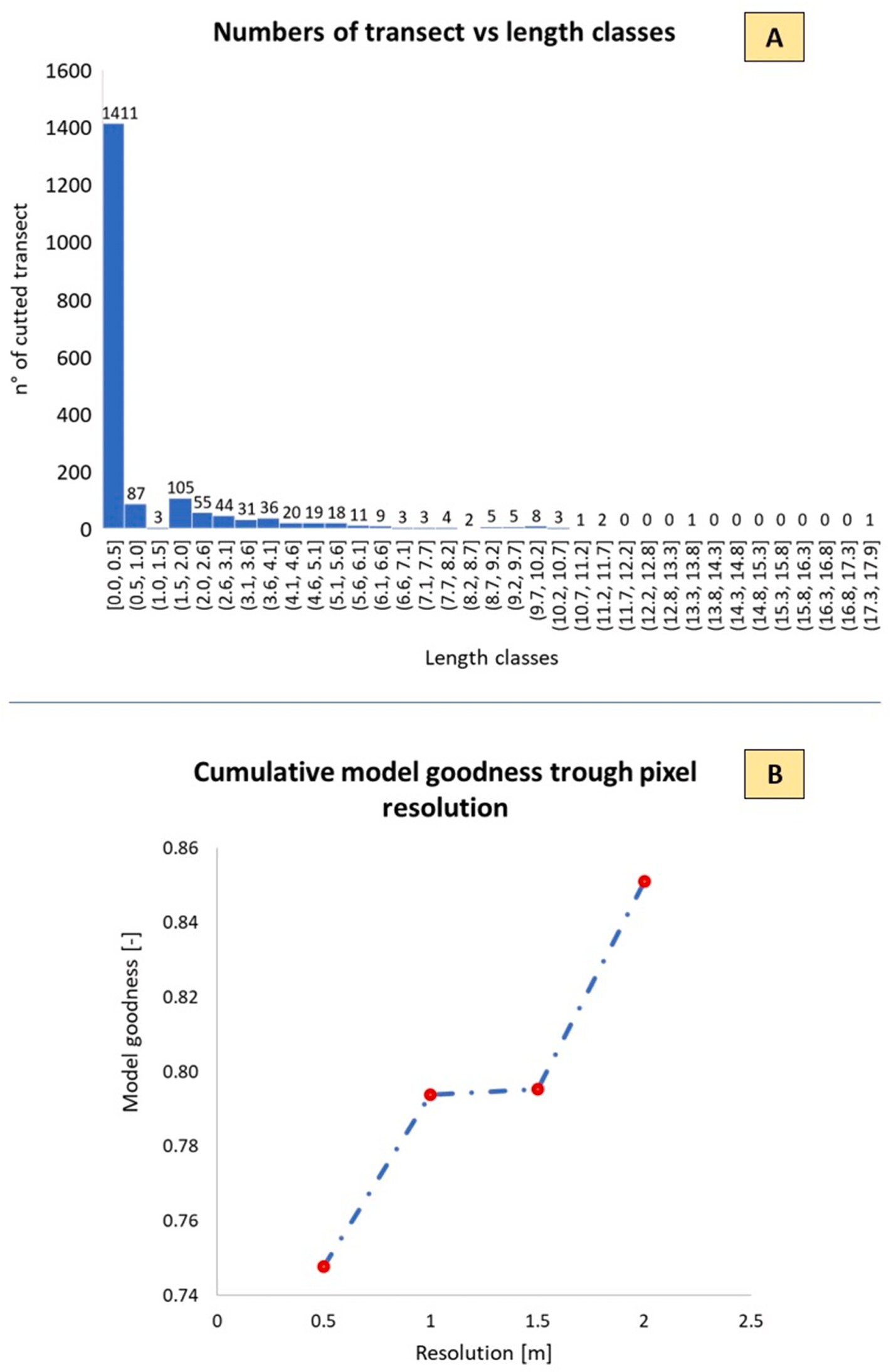
having a length greater than 1 m. The lengths are divided into classes with a variation of 0.50 m. The number of cut transects in the first two length classes of [Fig. 12](#_bookmark18) represent all cut transects with lengths less than the image resolution (can be understood as a surrogate for the IoU).

Specifically, the number of cut transects with length less than the pixel resolution of the image used was 1498 while transects with length greater than 1 m were found to be 389.

The percentage of cut transects that therefore identified a detected coastline deviation of less than pixel resolution from ground truth is 80



**Fig. 11.** Cut transects between the ground truth line and the detected line. Transects with length less than 1 m (in green) transects with length greater than 1 m (in red).



**Fig. 12.** In panel A is shown a histogram plot of transect lengths divided by length classes. In panel B is shown the trend in model goodness-of-fit as a function of the chosen geometric accuracy size.

percent. Consequently, the error percentage for all transects with length greater than 1 m is 20 %.

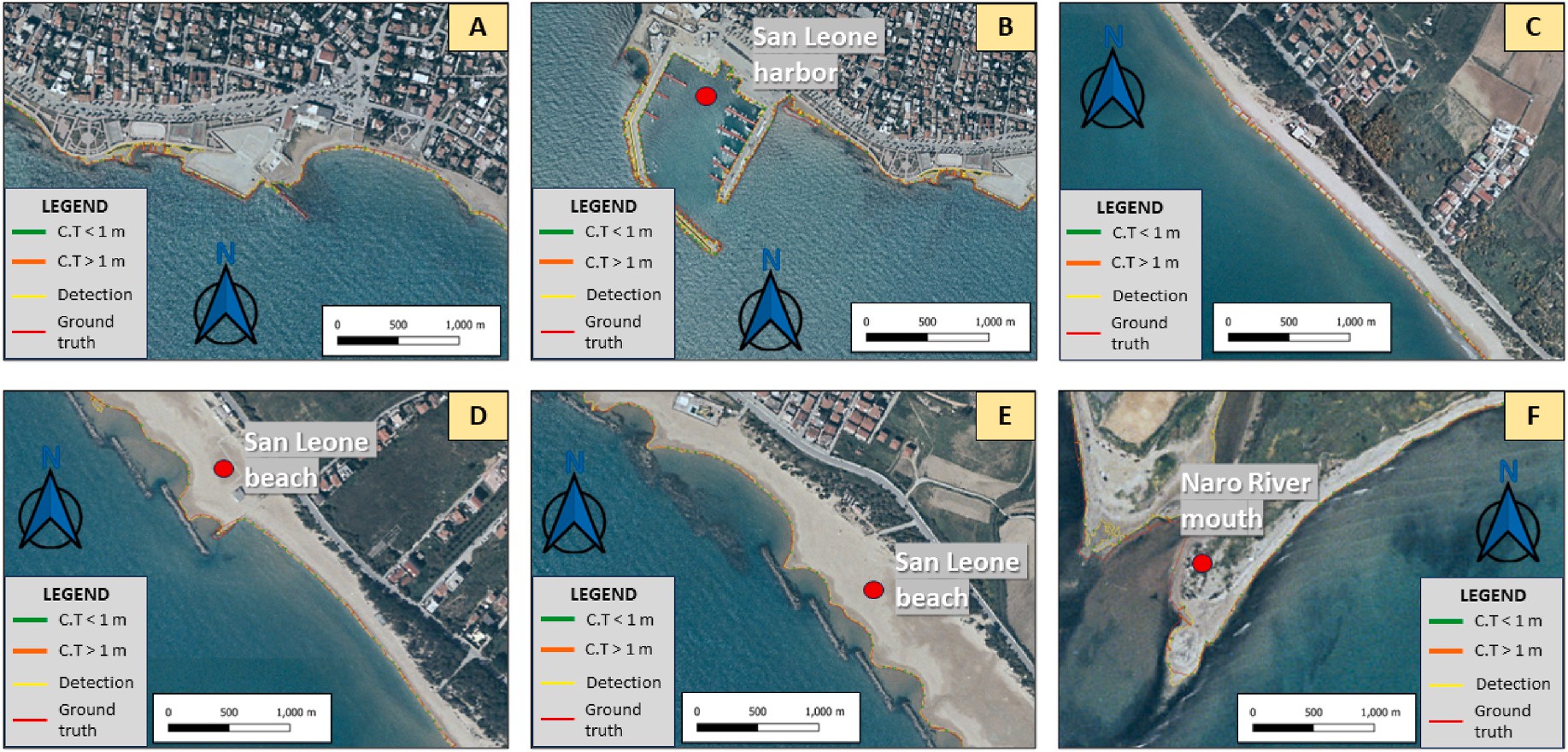
Considering that the geometric resolution of an image pixel is 1 m, it was assumed that the accuracy can be set equal to 2 times the geometric size of the pixel. Thus, cut transects with a length of less than 2 m represent those points where the coastline derived from the model has an uncertainty of at most 1 m with respect to the pixel where the real coastline is present.

In this case, the percentage of model error is 15% resulting in an increase in model goodness-of-fit of 85%. [Fig. 12](#_bookmark18), panel B shows the trend in model goodness-of-fit as a function of the chosen geometric

accuracy size (1 m–2 m).

These analyses confirm what was reported in section [4.1](#_bookmark11), suggesting how the model metrics of accuracy and IoU (for both testing and vali- dation) well represented the real behavior of the model, even with im- ages different from the starting dataset used for the testing and validation phase. In fact, considering transects with lengths less than 1 m, the goodness-of-fit value achieved is comparable to the IoU value while the accuracy is almost similar to the goodness-of-fit values considering even transects with lengths less than 2 m.

Details of the surveyed coastline and the ground truth datum are shown in [Fig. 13](#_bookmark19), where six zoom particular images are shown to



**Fig. 13.** Details of comparison between the coastline measured by the model (yellow line) and ground truth (red line). The first line (A to C) represents the details where the two lines deviate the most. The second line (D to F) represents the details where the two lines almost overlap.

represent the significantly differences (first row of [Fig. 13](#_bookmark19)) and simi- larities (second row) parts of the comparison.

In particular, the size of the transects turns out to be larger in some parts of the port area (A and B in [Fig. 13](#_bookmark19)). It should be noted that the model manages to faithfully reproduce the course of the boulders of the harbor breakwater reef unlike the ground truth line which is represented as a single straight line.

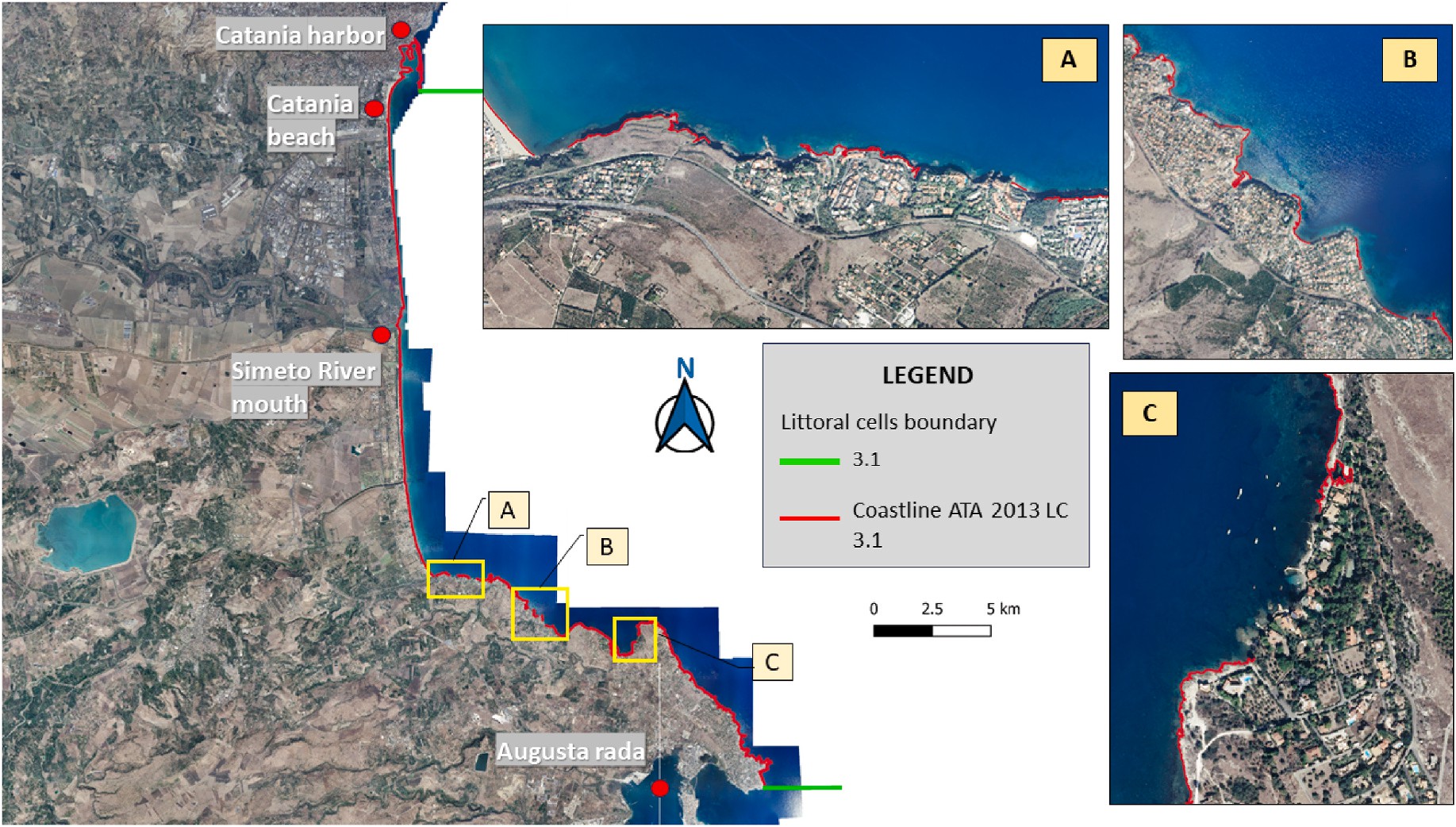
In this case the transect length is greater than 1 m suggesting therefore an error statistic of the model that might be lower when detection is done on images without harbors. In any case, since the coastline in the presence of ports is practically invariable and static over time, detection errors in these areas can be neglected. A zone of the study area in the presence of the foam produced by breaking waves is shown in panel C of [Fig. 13](#_bookmark19). In these areas the model was probably

unable to recognize these pixels as white-water because the breaking was captured practically attached to the coastline.

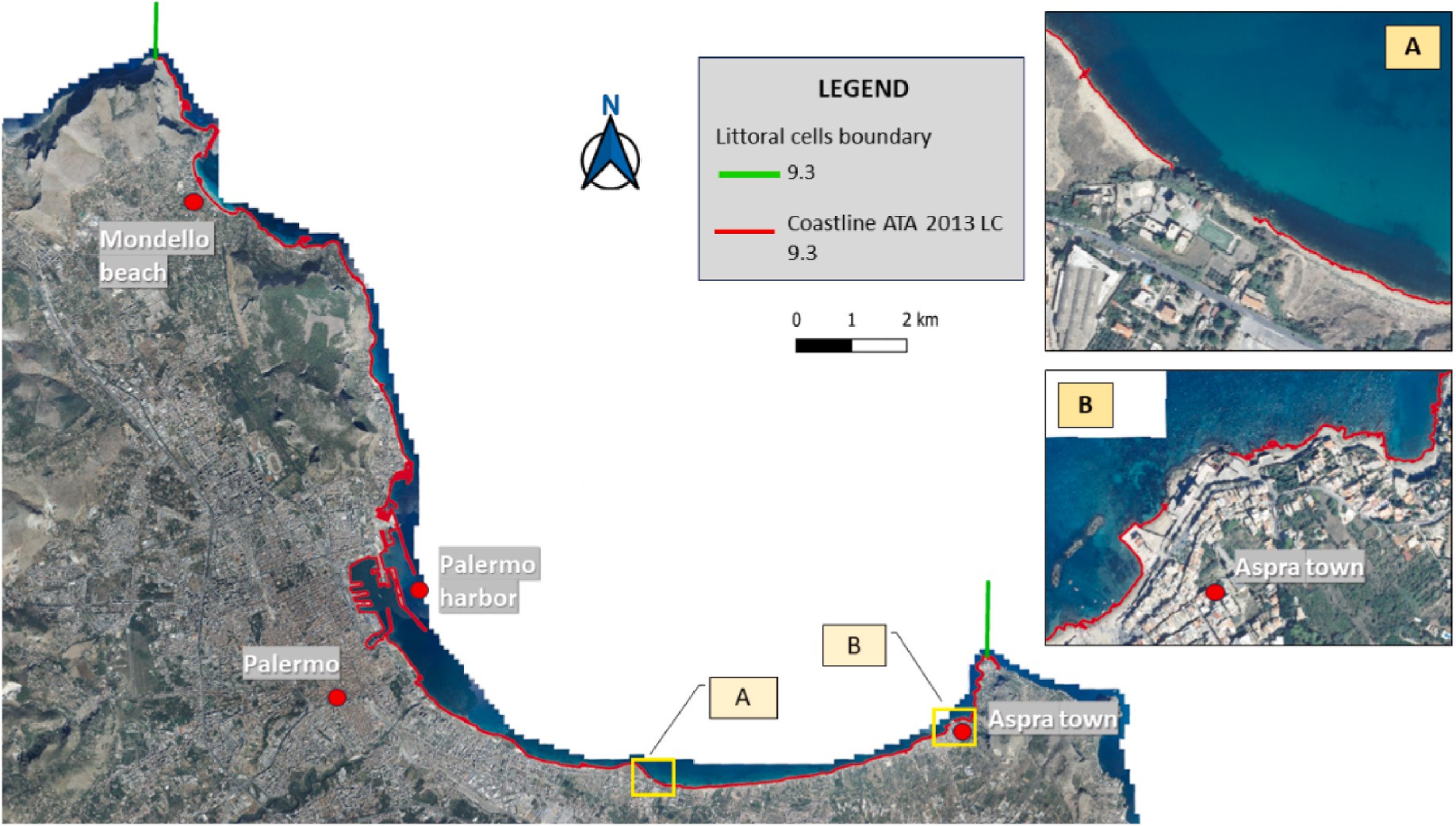
The second row of [Fig. 13](#_bookmark19)(D–F), on the other hand, shows some details in which the ground truth line and the line obtained from the model are almost identical. The transects in these areas therefore have the lowest length values. In these zones (particularly around the river mouth) the yellow line (detection model) seems to more naturally rep- resents and follows the coastline, identified as the transition line be- tween water and land.

* 1. *Coastline detection on second order Sicilian littoral cells*

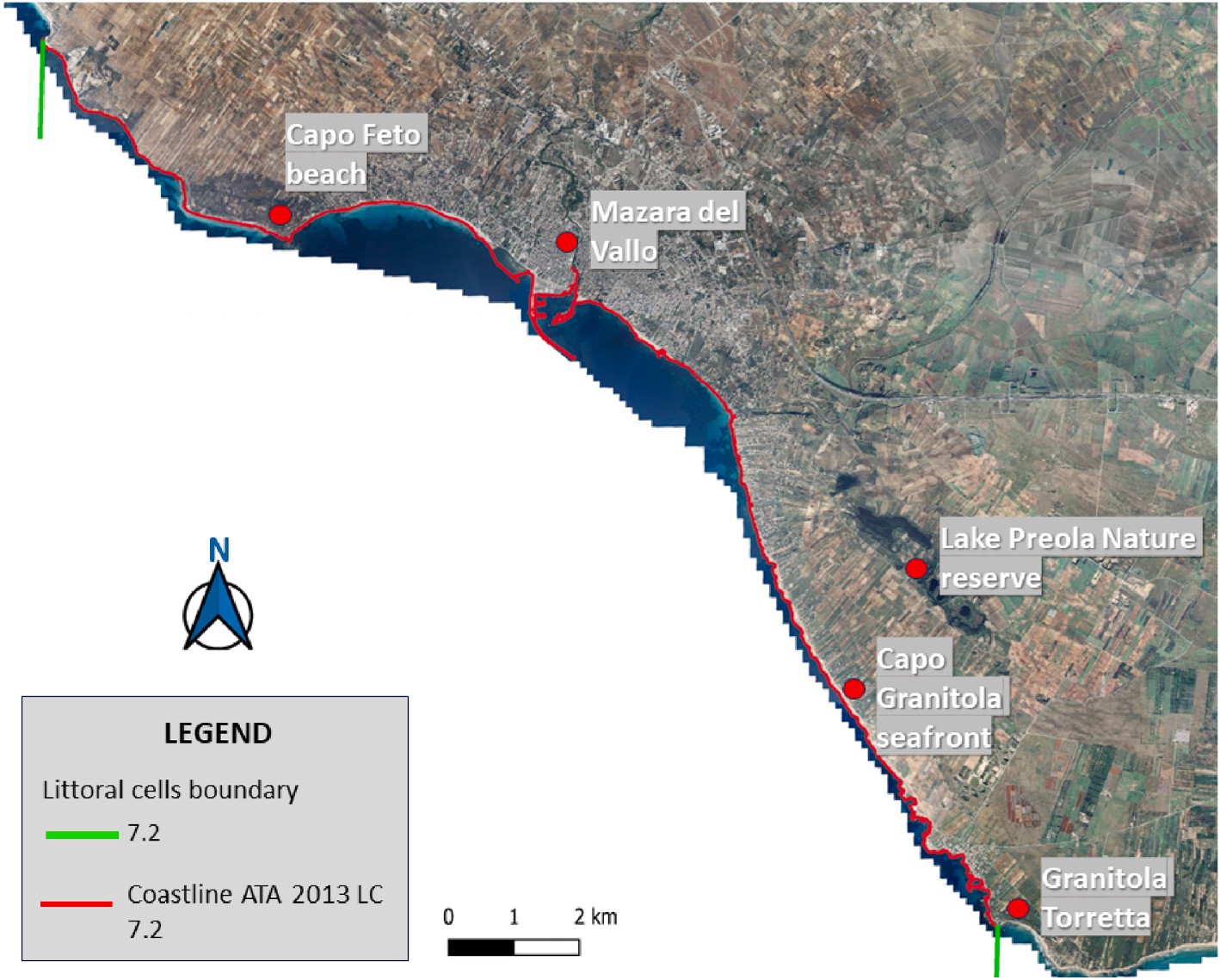
[Figs. 14–16](#_bookmark20) show the extracted coastlines with the red line for the Ionian, Tyrrhenian and Mediterranean coast, respectively. The proposed



**Fig. 14.** 3.1 Littoral cell coastline (showed in red). Panels A, B, C represent three areas where the model did not extract the coastline.



**Fig. 15.** 9.3 Littoral cell coastline (showed in red). Panels A, B, C represent three areas where the model did not extract the coastline.

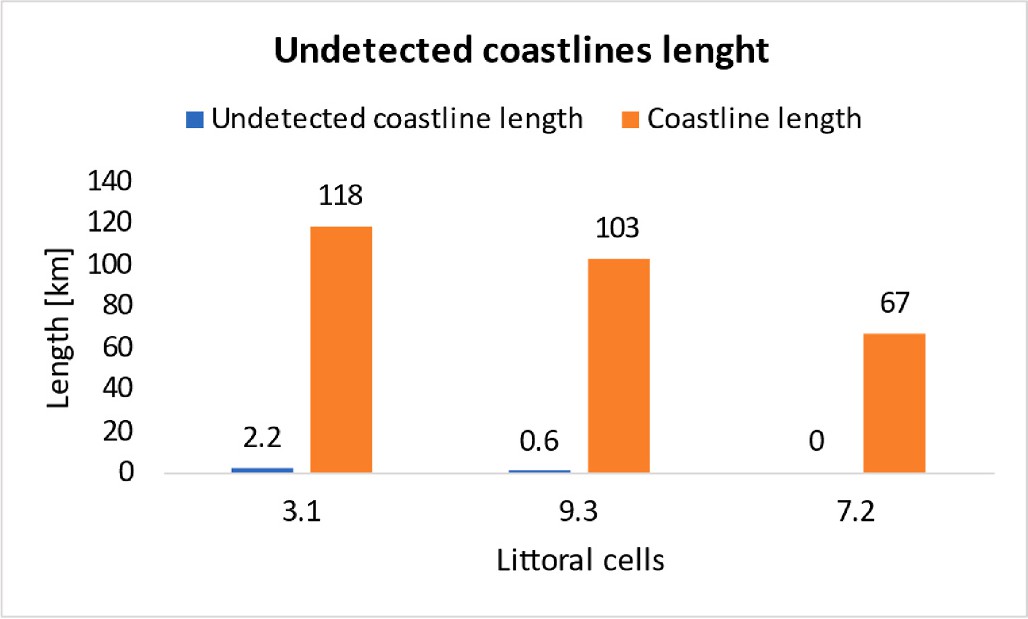


**Fig. 16.** 7.2 Littoral cell coastline (in red). In this case there are no areas where the model did not extract the coastline.

NN-based model effectively and successfully identified the coastline in images showing sandy beaches. The system also extracted the coastline in almost all features depicting cliffs, however, it had some shortcomings in extracting the coastline in areas where relatively darker pixels were sporadically present on the terrain. More in detail, the LC with a worse detection was 3.1, with 2.2 km of unidentified coastline compared with a total of 110 km. [Fig. 14](#_bookmark20) shows the coastline extracted from the model for LC 3.1.

In Panels A, B and C of the same Figure, zooms on some of the areas

where the coastline was not identified are plotted. In LC 9.3, the un- identified coastlines represent a total length of 0.6 km out of a total of 97 km. Again, [Fig. 15](#_bookmark21) shows the coastline extracted from the model for LC 9.3. In panels A, B of the same Figure, zooms are plotted on some of the areas where the coastline was not identified. The coastline related to LC 7.2 was fully extracted for a total of 64 km ([Fig. 16](#_bookmark22)). A histogram summarizing what has just been reported is shown in [Fig. 17](#_bookmark23).



**Fig. 17.** Comparison of coastline’s total length (orange bars) and total length of non-extracted coastline (blue bars) for each second order LC.

# Conclusions

A new approach to detect coastlines on regional scale studies was carried out. A CNN-based model for image segmentation was developed. This study demonstrates the use of satellite remote sensing as a fundamental tool for monitoring coastal changes while simultaneously addressing its inherent challenges. Using the U-net architecture with an input resolution of 512x512 pixels proved advantageous for training the model and generating accurate semantic segmentations for regional- scale image analysis. The model was improved using the recently pub- lished Coast Train dataset to ensure that it is based on the most up-to- date and especially specific coastal data. The multiclass segmentation capability of our model, which differs from traditional models, allows it to distinguish not only "water" and "land," but also the classes "built-up," "vegetation," bare land” and “beach”. This granularity provides a more

comprehensive perspective of the landscape.

Applying the model in post-processing to define coastlines as the boundary between "water" pixels and all other types shows its advanced coastline extraction capabilities. These sophisticated features ensure precise separation between different elements, resulting in more accu- rate coastline extraction. Performance metrics reveal an accuracy satu- ration point of 87 percent and an IoU close to 76 percent, reinforcing the quality of the segmentation achieved.

However, it is important to note some limitations. The model may not excel in high-resolution coastal studies (local scale, e.g. 1: 5.000, 1: 1000, etc.), and in some cases, it encountered difficulties in detecting coastlines in regions with darker pixels (e.g. shadows areas or seaweed). In the case of the image dataset used (satellite/airborne with respective ground truth classifications), although balanced, some pixel categories with similar bands (e.g., Dark Vegetation and Sea) might be subject to confusion by the classifier, especially if the input images intended for segmentation have already unbalanced RGB band classes of pixels. Such an occurrence may lead to model confusion, which may erroneously assign the two categories to the same group, generating higher error rates. Moreover, the coastal environment is known for its complexity, with variations in color, shape, and texture that can pose challenges even for an expert human eye. This complexity makes it challenging for the model to accurately identify and segment the various features of the environment. Additionally, overlapping of different land or cover classes in some areas of the image further complicates the model’s task of delineating precise boundaries between them, resulting

in partially correct or inaccurate segmentations.

On the other hand, when the classes are well balanced, the classifier works effectively ([Lee et al., 2022](#_bookmark55)), with a final accuracy of 87 %. Future studies will focus on further validating the model’s classification capa- bilities on all specified classes. To further improve the classification process and achieve even greater accuracy, we might consider

differentiated assignment of class weights during training. Specifically, implementing an additional *Keras* callback, “*class weight*”, during the training phase of the model. This callback will allow differentiated assignment of class weights during model loss calculation, considering the imbalance of classes in the dataset. Specifically, through the “*balanced*” function, the least represented classes will receive higher weights, while the most represented classes will receive lower weights. This will ensure that the model gives adequate attention to all classes during training, even if some classes are less frequent than others in the dataset.

Despite these difficulties, proposed model performance in the Sicil- ian case study underscores its adaptability even in highly anthropized areas. Tests conducted on various Sicilian coastal regions further consolidated its robustness, even in environments as diverse as sandy coasts and cliffs.

The efficiency of our model is not only theoretical; its high speed in generating coastlines has practical implications, enabling fast and effi- cient mapping and analysis of coastal areas. This speed makes it valuable for applications that require rapid assessment of coastal conditions, such as emergency response scenarios. In addition, its potential extends beyond simply surveying the current coastline; it opens the way for analysis of historical shoreline shifts, rather than analysis of land cover changes, offering valuable insights for coastal planning and management.

Our forward-looking approach is to further enhance the model’s capabilities, especially in terms of land use classification, ensuring that it remains at the forefront of coastline detection and geographic analysis.

# Code availability section

Name of the code/library: code name: SHODEM, library: TensorFlow 2.11, program run: Python 3.9.

Contact: +39 3349917567 e-mail: [pietro.scala@unipa.it](mailto:pietro.scala@unipa.it).

Hardware requirements: Windows 11 Pro operating system, a 12th Gen Intel® Core™ i7-12700 processor, 16 GB RAM DDR4 3200 MHz, Intel 770 UHD integrated graphics adapter, and a 1 TB nvme SSD.

Program language: Python.

The source codes are available for downloading at the link: <https://bitbucket.org/shodem/shodem/src/README/>

Stand-alone executable SHODEM model software upon first or cor- responding author request.

# CRediT authorship contribution statement

**Pietro Scala:** Conceptualization, Data curation, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Giorgio Manno:** Data curation, Investigation, Supervision, Validation, Visualization, Writing – review & editing, Methodology. **Giuseppe Ciraolo:** Investigation, Resources, Supervision, Validation, Writing – review & editing.

# Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

# Data availability

Data will be made available on request.

# Acknowledgments

This study was carried out within the RETURN Extended Partnership and received funding from the European Union Next-GenerationEU (National Recovery and Resilience Plan – NRRP, Mission 4,

Component 2, Investment 1.3 – D.D. 1243 2/8/2022, PE0000005).

# Appendix. section [1](#_bookmark0)

As described by [Seale et al. (2022)](#_bookmark74), the filters used in the analysis calculate the gradients on the x and y axes of an image labeled A, as shown in equations [(1) and (2](#_bookmark25)) below ([Seale et al., 2022](#_bookmark74)):

1 0 -1

⃒ ⃒

*Gx* = 2 0 -2 ∗ *A* (1)

⃒ 1 0 -1 ⃒

1 2 1

⃒ ⃒

*Gy* = 0 0 0 ∗ *A* (2)

⃒ -1 -2 -1 ⃒

In the Sobel edge detection method ([Vincent and Folorunso, 2009](#_bookmark82)), the gradient magnitude G is typically computed using the following equation (Equation [(3)](#_bookmark26)):

*G* = √*G*̅̅̅̅2̅̅̅+̅̅̅̅̅*G*̅̅̅2̅̅

*x*

*y*

(3)

The Sobel loss, comparing the target (p) with the prediction (p’), is defined as a loss function, with "i" representing a sample from the training image. The authors ([Seale et al., 2022](#_bookmark74)) provide the formula for this loss function, which is as follows:

*i*

ʹ 1 ∑ ( )2

*n*

*Sobel*(*p*, *p* ) = *n*

*i*=1

*Gpi* - *Gp*ʹ

(4)

# Appendix. section [2](#_bookmark2)

To encapsulate, while accuracy predominantly counts correct predictions, IoU delves deeper, analyzing the degree of overlap with true regions.

The mathematical expressions for these metrics are provided below (Equations [(5) and (6](#_bookmark27))):

*A* = *TP* + *TN*

*TP* + *TN* + *FP* + *FN*

*IoU* = *TP*

*TP* + *FP* + *FN*

(5)

(6)

Let TP represent the count of pixels accurately identified as part of a specific category, while TN indicates the count of pixels accurately determined not to be part of that category. FP stands for pixels wrongly categorized as part of the group, and FN signifies pixels wrongly left out of that group. [Fig. 3](#_bookmark6) shows a diagram illustrating the entire procedure of creating data and instructing the neural network model. By giving as input an image depicting a coastal environment, the pre-trained model provides as output a mask image (prediction image) of the same environment.

# Appendix. section [3](#_bookmark8)

We trained the proposed architecture using 256 x 256 as the input size and actually obtained lower metrics confirming what has been reported by (Lin et al., 312 2019; [Nekrasov et al., 2018](#_bookmark67); [Tanveer et al., 2022](#_bookmark79)). All other hyperparameters were left unchanged.

The two best training accuracy scores are 0.85 and 0.79 and training IoU 0.75 and 0.72 obtained from the 512 × 512 input (Kernel Size 3 × 3) and 256 × 256 input (Kernel Size 3 × 3) configuration, respectively. On the other hand, the two best validation accuracy scores are 0.85 and 0.80 while validation IoU of 0.72 and 0.68 obtained from the 512 × 512 (Kernel Size 3 × 3) input and 256 × 256 (Kernel Size 3 × 3) input configuration, respectively.

Note that the 256 × 256 metrics are worse than the results obtained by DeepLabV3 but if we consider that the computational training times of the latter are longer (about 24 h) and still comparable to those of Unet with input size of 512 x 512 then the U-Net architecture this input size appears to be the most suitable for coastal image segmentation.

# References

Aghdami-Nia, M., Shah-Hosseini, R., Rostami, A., Homayouni, S., 2022. Automatic coastline extraction through enhanced sea-land segmentation by modifying Standard U-Net. Int. J. Appl. Earth Obs. Geoinformation 109, 102785. [https://doi.org/](https://doi.org/10.1016/j.jag.2022.102785) [10.1016/j.jag.2022.102785](https://doi.org/10.1016/j.jag.2022.102785).

Albon, C., n.d. Machine Learning with Python Cookbook : Practical Solutions from Preprocessing to Deep Learning.

Anfuso, G., Martínez Del Pozo, J.A´., 2009. Assessment of coastal vulnerability through

the use of GIS tools in south sicily (Italy). Environ. Manag. 43, 533–545. [https://doi.](https://doi.org/10.1007/s00267-008-9238-8) [org/10.1007/s00267-008-9238-8](https://doi.org/10.1007/s00267-008-9238-8).

Anfuso, G., Martínez-del-Pozo, J.A´., Rangel-Buitrago, N., 2013. Morphological cells in

the ragusa littoral (sicily, Italy). J. Coast Conserv. 17, 369–377. [https://doi.org/](https://doi.org/10.1007/s11852-013-0233-8) [10.1007/s11852-013-0233-8](https://doi.org/10.1007/s11852-013-0233-8).

Anfuso, G., Martínez-del-Pozo, J.A´., Rangel-Buitrago, N., 2012. Bad practice in erosion management: the southern sicily case study. In: Cooper, J.A.G., Pilkey, O.H. (Eds.), Pitfalls of Shoreline Stabilization, Coastal Research Library. Springer, Netherlands, Dordrecht, pp. 215–233. <https://doi.org/10.1007/978-94-007-4123-2_13>.

Boak, E.H., Turner, I.L., 2005. Shoreline definition and detection: a review. J. Coast Res.

214, 688–703. <https://doi.org/10.2112/03-0071.1>.

Borzì, L., Anfuso, G., Manno, G., Distefano, S., Urso, S., Chiarella, D., Di Stefano, A., 2021. Shoreline evolution and environmental changes at the NW area of the Gulf of gela (sicily, Italy). Land 10, 1034. <https://doi.org/10.3390/land10101034>.

Brownlee, A.E.I., Adair, J., Haraldsson, S.O., Jabbo, J., 2021. Exploring the accuracy – energy trade-off in machine learning. In: 2021 IEEE/ACM International Workshop on Genetic Improvement (GI). Presented at the 2021 IEEE/ACM International Workshop on Genetic Improvement (GI). IEEE, Madrid, Spain, pp. 11–18. [https://](https://doi.org/10.1109/GI52543.2021.00011) [doi.org/10.1109/GI52543.2021.00011](https://doi.org/10.1109/GI52543.2021.00011).

Buscombe, D., Wernette, P., Fitzpatrick, S., Favela, J., Goldstein, E.B., Enwright, N.M., 2023. A 1.2 billion pixel human-labeled dataset for data-driven classification of coastal environments. Sci. Data 10, 46. [https://doi.org/10.1038/s41597-023-01929-](https://doi.org/10.1038/s41597-023-01929-2) [2](https://doi.org/10.1038/s41597-023-01929-2).

Buslaev, A., Iglovikov, V.I., Khvedchenya, E., Parinov, A., Druzhinin, M., Kalinin, A.A., 2020. Albumentations: fast and flexible image augmentations. Information 11, 125. <https://doi.org/10.3390/info11020125>.

Cao, W., Zhou, Y., Li, R., Li, X., 2020. Mapping changes in coastlines and tidal flats in developing islands using the full time series of Landsat images. Remote Sens.

Environ. 239, 111665 <https://doi.org/10.1016/j.rse.2020.111665>.

Castagno, K.A., Jim´enez-Robles, A.M., Donnelly, J.P., Wiberg, P.L., Fenster, M.S., Fagherazzi, S., 2018. Intense storms increase the stability of tidal bays. Geophys. Res. Lett. 45, 5491–5500. <https://doi.org/10.1029/2018GL078208>.

Chang, L., Chen, Y.-T., Wu, M.-C., Alkhaleefah, M., Chang, Y.-L., 2022. U-net for taiwan shoreline detection from SAR images. Rem. Sens. 14, 5135. [https://doi.org/](https://doi.org/10.3390/rs14205135) [10.3390/rs14205135](https://doi.org/10.3390/rs14205135).

[Chen, L.-C., Papandreou, G., Schroff, F., Adam, H., 2017. Rethinking atrous convolution](http://refhub.elsevier.com/S0098-3004(24)00187-0/sref14) [for semantic image segmentation. ArXiv Prepr. ArXiv170605587](http://refhub.elsevier.com/S0098-3004(24)00187-0/sref14).

Dang, K.B., Dang, V.B., Ngo, V.L., Vu, K.C., Nguyen, H., Nguyen, D.A., Nguyen, T.D.L., Pham, T.P.N., Giang, T.L., Nguyen, H.D., Hieu Do, T., 2022. Application of deep learning models to detect coastlines and shorelines. J. Environ. Manag. 320, 115732 <https://doi.org/10.1016/j.jenvman.2022.115732>.

Dhingra, N., Chogovadze, G., Kunz, A., 2021. Border-SegGCN: improving semantic segmentation by refining the border outline using graph convolutional network. In: 2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW). Presented at the 2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW). IEEE, Montreal, BC, Canada, pp. 865–875. [https://doi.org/](https://doi.org/10.1109/ICCVW54120.2021.00102) [10.1109/ICCVW54120.2021.00102](https://doi.org/10.1109/ICCVW54120.2021.00102).

Emanuel, K., 2005. Increasing destructiveness of tropical cyclones over the past 30 years.

Nature 436, 686–688. <https://doi.org/10.1038/nature03906>.

[Fernando, K.R.M., Tsokos, C.P., 2021. Dynamically weighted balanced loss: class](http://refhub.elsevier.com/S0098-3004(24)00187-0/sref18) [imbalanced learning and confidence calibration of deep neural networks. IEEE](http://refhub.elsevier.com/S0098-3004(24)00187-0/sref18) [Transact. Neural Networks Learn. Syst. 33, 2940–2951](http://refhub.elsevier.com/S0098-3004(24)00187-0/sref18).

[Galeone, P., 2019. Hands-On Neural Networks with TensorFlow 2.0: Understand](http://refhub.elsevier.com/S0098-3004(24)00187-0/sref19) [TensorFlow, from Static Graph to Eager Execution, and Design Neural Networks](http://refhub.elsevier.com/S0098-3004(24)00187-0/sref19).

[Girshick, R., Donahue, J., Darrell, T., Malik, J., 2014. Rich Feature Hierarchies for](http://refhub.elsevier.com/S0098-3004(24)00187-0/sref20) [Accurate Object Detection and Semantic Segmentation](http://refhub.elsevier.com/S0098-3004(24)00187-0/sref20).

[He, K., Gkioxari, G., Dolla´r, P., Girshick, R., 2018. Mask R-CNN](http://refhub.elsevier.com/S0098-3004(24)00187-0/sref22).

Hinkel, J., Lincke, D., Vafeidis, A.T., Perrette, M., Nicholls, R.J., Tol, R.S.J., Marzeion, B., Fettweis, X., Ionescu, C., Levermann, A., 2014. Coastal flood damage and adaptation costs under 21st century sea-level rise. Proc. Natl. Acad. Sci. USA 111, 3292–3297. <https://doi.org/10.1073/pnas.1222469111>.

Intergovernmental Panel On Climate Change (Ipcc), 2022. The Ocean and Cryosphere in a Changing Climate: Special Report of the Intergovernmental Panel on Climate Change, first ed. Cambridge University Press. [https://doi.org/10.1017/](https://doi.org/10.1017/9781009157964) [9781009157964](https://doi.org/10.1017/9781009157964).

[Kandel, I., Castelli, M., 2020. The effect of batch size on the generalizability of the](http://refhub.elsevier.com/S0098-3004(24)00187-0/sref25) [convolutional neural networks on a histopathology dataset. ICT Express 6, 312–315](http://refhub.elsevier.com/S0098-3004(24)00187-0/sref25).

Kattenborn, T., Leitloff, J., Schiefer, F., Hinz, S., 2021. Review on convolutional neural networks (CNN) in vegetation remote sensing. ISPRS J. Photogrammetry Remote Sens. 173, 24–49. <https://doi.org/10.1016/j.isprsjprs.2020.12.010>.

Kopp, R.E., Horton, R.M., Little, C.M., Mitrovica, J.X., Oppenheimer, M., Rasmussen, D. J., Strauss, B.H., Tebaldi, C., 2014. Probabilistic 21st and 22nd century sea-level projections at a global network of tide-gauge sites. Earth’s Future 2, 383–406. <https://doi.org/10.1002/2014EF000239>.

[Lee, Y., Sim, W., Park, J., Lee, J., 2022. Evaluation of hyperparameter combinations of](http://refhub.elsevier.com/S0098-3004(24)00187-0/sref29) [the U-net model for land cover classification. Forests 13, 1813](http://refhub.elsevier.com/S0098-3004(24)00187-0/sref29).

Lin, G., Milan, A., Shen, C., Reid, I., 2017. RefineNet: multi-path refinement networks for high-resolution semantic segmentation. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Presented at the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, Honolulu, HI,

pp. 5168–5177. <https://doi.org/10.1109/CVPR.2017.549>.

Lin, K., Gong, L., Huang, Y., Liu, C., Pan, J., 2019. Deep learning-based segmentation and quantification of cucumber powdery mildew using convolutional neural network. Front. Plant Sci. 10, 155. <https://doi.org/10.3389/fpls.2019.00155>.

Lionello, P., O¨ zsoy, E., Planton, S., Zanchetta, G., 2017. Climate variability and change in

the mediterranean region. Global Planet. Change 151, 1–3. [https://doi.org/](https://doi.org/10.1016/j.gloplacha.2017.04.005) [10.1016/j.gloplacha.2017.04.005](https://doi.org/10.1016/j.gloplacha.2017.04.005).

[Liu, H., Jezek, K.C., 2004. Automated Extraction of Coastline from Satellite Imagery by](http://refhub.elsevier.com/S0098-3004(24)00187-0/sref33) [Integrating Canny Edge Detection and Locally Adaptive Thresholding Methods](http://refhub.elsevier.com/S0098-3004(24)00187-0/sref33).

Liu, Y., Wang, X., Ling, F., Xu, S., Wang, C., 2017. Analysis of coastline extraction from landsat-8 OLI imagery. Water 9, 816. <https://doi.org/10.3390/w9110816>.

Manno, G., Azzara, G., Lo Re, C., Martinello, C., Basile, M., Rotigliano, E., Ciraolo, G., 2022a. An approach for the validation of a coastal erosion vulnerability index: an application in sicily. J. Mar. Sci. Eng. 11, 23. [https://doi.org/10.3390/](https://doi.org/10.3390/jmse11010023) [jmse11010023](https://doi.org/10.3390/jmse11010023).

[Manno, G., Ciraolo, G., 2015. Diachronic analysis of the shoreline in san Leone beach](http://refhub.elsevier.com/S0098-3004(24)00187-0/sref36) [(Agrigento-Sicily). In: Establishment of an Integrated Italy-Malta Cross-Border](http://refhub.elsevier.com/S0098-3004(24)00187-0/sref36) [System of Civil Protection-Geological Aspects. Aracne, pp. 117–136](http://refhub.elsevier.com/S0098-3004(24)00187-0/sref36).

Manno, G., Lo Re, C., Basile, M., Ciraolo, G., 2022b. A new shoreline change assessment approach for erosion management strategies. Ocean Coast Manag. 225, 106226 <https://doi.org/10.1016/j.ocecoaman.2022.106226>.

Molina, R., Manno, G., Lo Re, C., Anfuso, G., Ciraolo, G., 2020. A methodological approach to determine sound response modalities to coastal erosion processes in mediterranean andalusia (Spain). J. Mar. Sci. Eng. 8, 154. [https://doi.org/10.3390/](https://doi.org/10.3390/jmse8030154) [jmse8030154](https://doi.org/10.3390/jmse8030154).

Muis, S., Apecechea, M.I., Dullaart, J., de Lima Rego, J., Madsen, K.S., Su, J., Yan, K., Verlaan, M., 2020. A high-resolution global dataset of extreme sea levels, tides, and storm surges, including future projections. Front. Mar. Sci. 7, 263. [https://doi.org/](https://doi.org/10.3389/fmars.2020.00263) [10.3389/fmars.2020.00263](https://doi.org/10.3389/fmars.2020.00263).

Murakami, H., Vecchi, G.A., Underwood, S., 2017. Increasing frequency of extremely severe cyclonic storms over the Arabian Sea. Nat. Clim. Change 7, 885–889. [https://](https://doi.org/10.1038/s41558-017-0008-6) [doi.org/10.1038/s41558-017-0008-6](https://doi.org/10.1038/s41558-017-0008-6).

[Nekrasov, V., Shen, C., Reid, I., 2018. Light-Weight RefineNet for Real-Time Semantic](http://refhub.elsevier.com/S0098-3004(24)00187-0/sref41) [Segmentation](http://refhub.elsevier.com/S0098-3004(24)00187-0/sref41).

Paravolidakis, V., Ragia, L., Moirogiorgou, K., Zervakis, M., 2018. Automatic coastline extraction using edge detection and optimization procedures. Geosciences 8, 407. <https://doi.org/10.3390/geosciences8110407>.

Pollard, J.A., Spencer, T., Brooks, S.M., Christie, E.K., Mo¨ller, I., 2020. Understanding spatio-temporal barrier dynamics through the use of multiple shoreline proxies. Geomorphology 354, 107058. <https://doi.org/10.1016/j.geomorph.2020.107058>.

[Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional Networks for](http://refhub.elsevier.com/S0098-3004(24)00187-0/sref44) [Biomedical Image Segmentation](http://refhub.elsevier.com/S0098-3004(24)00187-0/sref44).

Sapkota, Y., White, J.R., 2019. Marsh edge erosion and associated carbon dynamics in coastal Louisiana: a proxy for future wetland-dominated coastlines world-wide.

Estuar. Coast Shelf Sci. 226, 106289 <https://doi.org/10.1016/j.ecss.2019.106289>.

Satta, A., Puddu, M., Venturini, S., Giupponi, C., 2017. Assessment of coastal risks to climate change related impacts at the regional scale: the case of the Mediterranean region. Int. J. Disaster Risk Reduc. 24, 284–296. [https://doi.org/10.1016/j.](https://doi.org/10.1016/j.ijdrr.2017.06.018) [ijdrr.2017.06.018](https://doi.org/10.1016/j.ijdrr.2017.06.018).

Scardino, G., Scicchitano, G., Chirivì, M., Costa, P.J.M., Luparelli, A., Mastronuzzi, G., 2022. Convolutional neural network and optical flow for the assessment of wave and tide parameters from video analysis (LEUCOTEA): an innovative tool for coastal monitoring. Rem. Sens. 14, 2994. <https://doi.org/10.3390/rs14132994>.

Seale, C., Redfern, T., Chatfield, P., Luo, C., Dempsey, K., 2022. Coastline detection in satellite imagery: a deep learning approach on new benchmark data. Remote Sens. Environ. 278, 113044 <https://doi.org/10.1016/j.rse.2022.113044>.

Shi, J., Dang, J., Cui, M., Zuo, R., Shimizu, K., Tsunoda, A., Suzuki, Y., 2021.

Improvement of damage segmentation based on pixel-level data balance using VGG- unet. Appl. Sci. 11, 518. <https://doi.org/10.3390/app11020518>.

[Stocker, T., Qin, D. (Eds.), 2013. Climate Change 2013: the Physical Science Basis:](http://refhub.elsevier.com/S0098-3004(24)00187-0/sref50) [Summary for Policymakers, a Report of Working Group I of the IPCC: Technical](http://refhub.elsevier.com/S0098-3004(24)00187-0/sref50) [Summary, a Report Accepted by Working Group I of the IPCC but Not Approved in](http://refhub.elsevier.com/S0098-3004(24)00187-0/sref50) [Detail: and Frequently Asked Questions: Part of the Working Group I Contribution to](http://refhub.elsevier.com/S0098-3004(24)00187-0/sref50) [the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.](http://refhub.elsevier.com/S0098-3004(24)00187-0/sref50)

[WMO, UNEP, s.l](http://refhub.elsevier.com/S0098-3004(24)00187-0/sref50).

Sultan, A.S., Elgharib, M.A., Tavares, T., Jessri, M., Basile, J.R., 2020. The use of artificial intelligence, machine learning and deep learning in oncologic histopathology.

J. Oral Pathol. Med. 49, 849–856. <https://doi.org/10.1111/jop.13042>. Syvitski, J.P.M., Kettner, A.J., Overeem, I., Hutton, E.W.H., Hannon, M.T.,

Brakenridge, G.R., Day, J., Vo¨ro¨smarty, C., Saito, Y., Giosan, L., Nicholls, R.J., 2009.

Sinking deltas due to human activities. Nat. Geosci. 2, 681–686. [https://doi.org/](https://doi.org/10.1038/ngeo629) [10.1038/ngeo629](https://doi.org/10.1038/ngeo629).

Tanveer, M., Kim, B., Hong, J., Sim, S.-H., Cho, S., 2022. Comparative study of lightweight deep semantic segmentation models for concrete damage detection. Appl. Sci. 12, 12786 <https://doi.org/10.3390/app122412786>.

Toure, S., Diop, O., Kpalma, K., Maiga, A., 2019. Shoreline detection using optical remote sensing: a review. ISPRS Int. J. Geo-Inf. 8, 75. <https://doi.org/10.3390/ijgi8020075>.

[Tyagi, S., Mittal, S., 2020. Sampling approaches for imbalanced data classification](http://refhub.elsevier.com/S0098-3004(24)00187-0/sref56) [problem in machine learning. Presented at the Proceedings of ICRIC 2019: Recent](http://refhub.elsevier.com/S0098-3004(24)00187-0/sref56) [Innovations in Computing. Springer, pp. 209–221](http://refhub.elsevier.com/S0098-3004(24)00187-0/sref56).

Vincent, O., Folorunso, O., 2009. A descriptive algorithm for Sobel image edge detection.

Presented at the InSITE 2009: Informing Science + IT Education Conference. <https://doi.org/10.28945/3351>.

Vitousek, S., Barnard, P.L., Fletcher, C.H., Frazer, N., Erikson, L., Storlazzi, C.D., 2017. Doubling of coastal flooding frequency within decades due to sea-level rise. Sci. Rep. 7, 1399. <https://doi.org/10.1038/s41598-017-01362-7>.

Walsh, K.J.E., McBride, J.L., Klotzbach, P.J., Balachandran, S., Camargo, S.J.,

Holland, G., Knutson, T.R., Kossin, J.P., Lee, T., Sobel, A., Sugi, M., 2016. Tropical cyclones and climate change. WIREs Clim. Change 7, 65–89. [https://doi.org/](https://doi.org/10.1002/wcc.371) [10.1002/wcc.371](https://doi.org/10.1002/wcc.371).

[Wernette, P.A., Buscombe, D.D., Favela, J., Fitzpatrick, S., Goldstein, E., 2022. Coast](http://refhub.elsevier.com/S0098-3004(24)00187-0/sref60) [Train–Labeled imagery for training and evaluation of data-driven models for image](http://refhub.elsevier.com/S0098-3004(24)00187-0/sref60) [segmentation. U.S. Geological Survey data release](http://refhub.elsevier.com/S0098-3004(24)00187-0/sref60).

Wu, L., Tajima, Y., Yamanaka, Y., Shimozono, T., Sato, S., 2019. Study on characteristics of SAR imagery around the coast for shoreline detection. Coast Eng. J. 61, 152–170. <https://doi.org/10.1080/21664250.2018.1560685>.

Zhu, Q., Li, P., Li, Z., Pu, S., Wu, X., Bi, N., Wang, H., 2021. Spatiotemporal changes of coastline over the yellow river delta in the previous 40 Years with optical and SAR remote sensing. Rem. Sens. 13, 1940. <https://doi.org/10.3390/rs13101940>.

Zou, X., Xiao, F., Yu, Z., Li, Y., Lee, Y.J., 2023. Delving deeper into anti-aliasing in ConvNets. Int. J. Comput. Vis. 131, 67–81. [https://doi.org/10.1007/s11263-022-](https://doi.org/10.1007/s11263-022-01672-y) [01672-y](https://doi.org/10.1007/s11263-022-01672-y).