Automatic coastline extraction through enhanced sea-land segmentation by modifying Standard U-Net

Mohammad Aghdami-Nia [a], Reza Shah-Hosseini [a,*], Amirhossein Rostami [a], Saeid Homayouni [a,b]

[a] School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, Tehran, Iran
[b] Institut national de la recherche scientifique, Centre Eau Terre Environnement, Québec City, QC, Canada

ARTICLE INFO

ABSTRACT

Sea-land segmentation (SLS) is an essential remote sensing task for various coastal and environmental studies such as coastline extraction, coastal erosion, coastal area monitoring, and ship or iceberg detection. This study aims at improving the SLS performance by modifying the Standard U-Net (SUN) model and developing an automatic coastline extraction framework. SUN generally has an acceptable performance in many applications. However, better SLS outputs are needed for reliable coastline extraction. In our proposed framework, we firstly analyzed three different input images, including Red-Green-Blue (RGB), Normalized Difference Water Index (NDWI), and Near-Infrared (NIR) images. Secondly, we modified the SUN architecture to improve the segmentation results. The main modifications are using different loss functions and two fusion methods for RGB and NIR images. The segmentation results were then passed into the subsequent automatic coastline extraction pipeline based on morphological operations and pixel connectivity analysis. The training and testing steps were accomplished utilizing a benchmark dataset of China's coastal areas. Moreover, another dataset consisting of a time series of Landsat-8 imagery from the southern Caspian Sea coastlines was collected to evaluate coastline extraction efficiency. The results indicate that the proposed modifications could effectively enhance the performance of the SUN, with the most significant improvement to the Intersection over Union (IoU) score being as high as 1.68% and 8.95% in China and Caspian Sea datasets, respectively, while outperforming other state-of-the-art models including FC-DenseNet and DeepLabV3+.

1. Introduction

Sea-land segmentation (SLS) aims at separating and mapping the sea and land regions from remote sensing images of coastal areas. These maps are a valuable resource for coastal areas and ecosystems studies. However, obtaining these maps from optical remote sensing images is quite challenging due to factors like intensity inhomogeneity, complex land texture, and low contrast at the sea-land interface, making it spectrally difficult to discern between the sea and land regions (Elkhateeb et al., 2021). Many applications such as coastal management (Hamylton and Prosper, 2012), water resources management (Giardino et al., 2010), and coastline detection (Qiao et al., 2018) can benefit from these maps. Another application of these maps is in ship detection (Wu et al., 2009), where it is considered a major preprocessing step in which the land region is excluded from the search area. Improving the methodologies used for SLS is of great importance, and many researchers constantly offer new contributions to this topic.

The goal of semantic segmentation is to classify individual pixels in a given image into a specific number of classes, which in the case of SLS are sea and land classes. Historically, SLS was carried out by manually delineating satellite images, which is immensely time-consuming and labor-intensive. However, many semi-automatic and automatic approaches have been proposed for the SLS of optical remote sensing images in recent years that are not based on deep learning (DL). Feyisa et al. (2014) suggested a new Automated Water Extraction Index (AWEI) for automatic binary classification of water bodies and dryland areas that relied on thresholding. Evaluations on Landsat 5 TM images revealed that AWEI improved the classification accuracy compared to Modified Normalized Difference Water Index (MNDWI) and Maximum Likelihood methods. Based on the texture of sea and land surfaces, Huihui et al. (2016) developed the Gray Smoothness Ratio (GSR) descriptor for SLS of infrared images. Support vector machine (SVM) used the GSR to segment the images roughly, and the Otsu algorithm was adopted to segment sea-land boundaries. This method could reduce
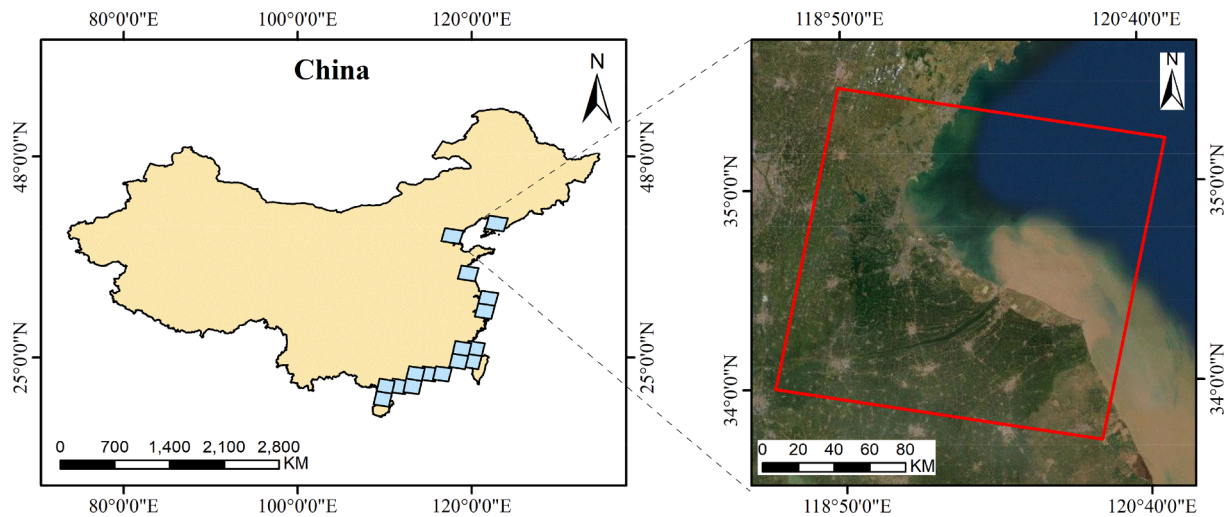
Fig. 1. China dataset study area. Blue squares represent the selected Landsat-8 images and the red square in the large-scale image corresponds to the image with Path 120 and Row 036. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

the computation complexity while offering a better segmentation result than the relevant approaches. Liu et al. (2017) presented a novel coarse-to-fine model for the SLS of high-resolution panchromatic imagery by merging the improved multi-scale normalized cut (IMNcut) and improved Chan-Vese methods. Assessments on different images revealed the better performance of this model compared to a few other models like local binary pattern (LBP). Lei et al. (2018) proposed a new method for segmenting infrared images using superpixels and multi-scale features. This model outperformed traditional algorithms when evaluations were carried out using Landsat-8 infrared images.

Recent advances in DL methods have led to their implementation in many remote sensing fields such as active fire detection (Rostami et al., 2022), building segmentation (Khoshboresh-Masouleh et al., 2020), and hyperspectral image classification (Ansari et al., 2021) where they have outperformed many conventional methods. This revolution has caused a shift in image processing and artificial intelligence from machine learning approaches such as SVM (Ranjbar et al., 2021) and random forest (RF) (Zarei et al., 2021) towards DL in the remote sensing community since 2014 (Ma et al., 2019). CNNs are the most sought-after and published DL networks in the remote sensing community (Ball et al., 2017). In recent years, these networks have achieved remarkable improvements (Zhang et al., 2021). There have been multiple cases of applying DL methods, especially CNNs, to the SLS problem.

Cheng et al. (2017) suggested the CNN model FusionNet to segment harbor images in high-resolution Google Earth images. Their model performed SLS and ship detection simultaneously and outperformed other models like SegNet. Li et al. (2021a) suggested an automatic framework for mapping Germany's national surface water by training the ResNet model using OpenStreetMap (OSM) and Sentinel-2 data. This model obtained an accuracy score of 86.32% when comparing their resulting open water products to existing water products. Li et al. (2018) proposed DeepUNet for the segmentation of high-resolution optical images. They replaced the convolutions with DownBlocks and UpBlocks, which improved the accuracy by 1–2% compared to other models like U-Net. Shamsolmoali et al. (2019), utilizing residual blocks that were densely connected alongside the convolution layers, developed RDU-Net. Evaluations on images from Google Earth and ISPRS_Benchmark showed that this model outperformed other methods like DenseNet.

Li et al. (2021b) introduced a pixel-based CNN for water segmentation using region of interest (ROI) labels for training instead of full-label datasets. Assessments with Landsat-8 images revealed this method's better performance than other methods like thresholding Normalized Difference Water Index (NDWI). Yang et al. (2020b) collected a sea-land benchmark dataset from Landsat-8 images and assessed the SLS

performance of different state-of-the-art architectures including RefineNet, FC-DenseNet, DeepLabV3+, PSPNet, SegNet, and U-Net.

Their results revealed that DeepLabV3+ and FC-DenseNet achieved the best performance with the latter having the highest Intersection over Union (IoU) score. Cui et al. (2021) suggested the scale-adaptive network SANet that substitutes the serial convolution operation for a multi-scale feature learning module. Using the Gaofen-1 images for assessment revealed that this model improved the accuracy compared to other models like DeepUNet while achieving sharper boundaries. Although many DL-based SLS frameworks have been suggested in various studies, most of them have not investigated the optimization of the performance of an individual model regarding the variables such as input data type and model structure. Also, these studies have mainly focused on comparing different state-of-the-art methods instead of further enhancing an efficient model to reach the highest possible accuracy.

Another topic that is closely related to SLS is coastline extraction. Coastline extraction is the process of detecting the pixels pertaining to the sea-land interface, which is fundamentally similar to edge detection. There are multiple methods for automatic and semi-automatic coastline extraction in the literature. Van der Werff (2019) proposed a supervised edge detection method for mapping several shoreline indicators based on the analysis of differences in spectral albedo caused by the moisture content of the sand. Comparisons with two other indicators derived from a field survey and a digital elevation model revealed that the positional error in this method was within the acceptable margin. Vos et al. (2019) evaluated the potential of publicly available optical satellite imagery for detecting the variability in coastline position based on Multilayer Perceptron classification and sub-pixel border segmentation. The proposed method was tested on five sites with long-term in situ measurements. The root mean square error (RMSE) of the extracted coastlines in these test sites ranged from 7.3 to 12.7 m.

Using GPS survey data, Kelly and Gontz (2018) evaluated all known Landsat-8 water indices for the automatic mapping of coastlines based on a thresholding approach. Amongst the seven indices, MNDWI achieved the best performance with an accuracy of 88.4%. San and Ulusar (2018) introduced the spatial uncertainty algorithm SLiP-SUM for the semi-automatic detection of shorelines based on the snake algorithm. They tested the suggested algorithm on a dynamic coastal area and predicted the future shorelines for 2020, 2025, 2030, and 2035. Rogers et al. (Rogers et al., 2021) developed the automatic coastal vegetation line detection tool VEdge_Detector using a DL method based on holistically nested edge detection architecture. Assessments on high spatial resolution optical images revealed that the optimum input spectral band
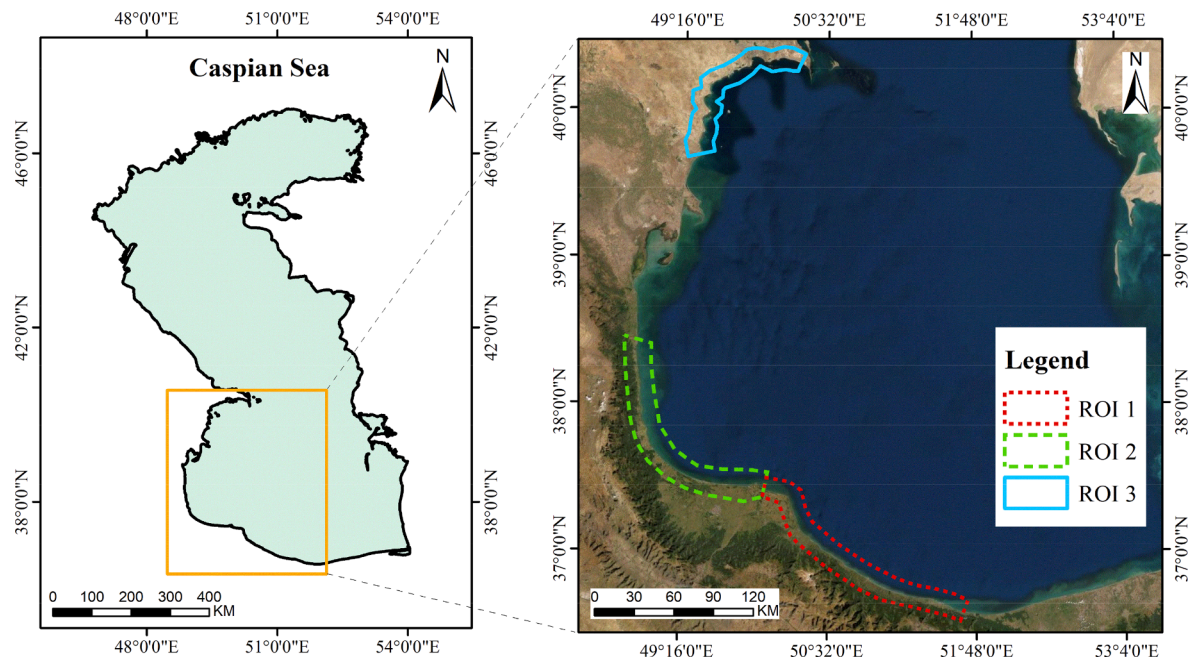
Fig. 2. Caspian Sea study area. The dotted red, dashed green, and solid blue polygons represent ROIs 1, 2, and 3 respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

combination was Red-Green-NIR . The mean distance error of the outputs was less than 6 m (2 pixels) compared to positional measurements and manually digitized vegetation lines.

Despite all the contributions in the literature regarding SLS with DL methods and automatic coastline extraction, little attention has been paid to merging these two topics. On the other hand, only a few studies have experimented with different band combinations to find the optimum input type for SLS in DL methods, and further assessments are needed in the community. Moreover, most related studies have only focused on comparing different DL models instead of further enhancing an optimum model to reach the highest possible accuracy. Considering the abovementioned issues, this study aims to propose an enhanced SLS and automatic coastline extraction framework by the following contributions:

– Modifications to the SUN architecture are introduced to increase segmentation accuracy. These modifications include using the Jaccard loss function, implementing a double encoder architecture for the late fusion of Red-Green-Blue (RGB) and Near-Infrared (NIR) images, and utilizing a residual operation in convolution blocks.
– NDWI, NIR, and RGB images are used as the input to designed models, and the results are compared. To our best knowledge, this is the first study to utilize the NDWI images as the input for SLS in DL methods.
– A post-processing pipeline is suggested to automatically extract coastlines from the binary segmentation outputs of the trained models based on pixel connectivity classification and morphological operations.

2. Remote sensing datasets

2.1. Datasets acquisition

2.1.1. China dataset

The primary dataset in this study is based on the benchmark SLS dataset collected by Yang et al. (2020b). They chose the offshore coastal area of China as their study area (Fig. 1). From all the Landsat-8 OLI image collections of the study area, 17 and 12 unique full-frame images with less than 5% cloud cover were selected for training and testing purposes, respectively. The Landsat-8 full image size was too large to be fed directly into the model, so they were split into 512 × 512 pixel patches. At the end of the patching process, 1,950 training samples and 1,411 testing samples were generated. The ground-truth images corresponding to each patch were manually annotated with LabelMe (Russell et al., 2008). Finally, Red-Green-Blue and NIR-SWIR1-Red spectral band combinations were selected separately in each sample to create two subdatasets. These two sub-datasets were used to train two instances of the same models. It should be noted that the samples in the China dataset were of PNG data type, without any spatial information.

2.1.2. Caspian Sea dataset

In this study, a second dataset has been collected from the Caspian Sea southern coastlines for the sake of further testing the models that were trained on the China benchmark dataset. This dataset has also been used to extract the southern Caspian Sea coastlines. The Caspian Sea is the most extensive closed water body in the world without tides presence, and its water level repeatedly oscillates (Firoozfar et al., 2012). Therefore, continually monitoring its coastlines is of great importance. Tides introduce changes in the position of coastlines, making tidal level observations necessary for correcting their effects (Chen and Chang, 2009). Consequently, the lack of tides in the Caspian Sea is an asset for coastline extraction.

Three ROIs have been selected in the study area so that each one would not exceed a single Landsat-8 frame area (Fig. 2). While selecting these ROIs, two coastal regions with simple and complex morphologically and geometrically features have been considered. ROIs 1 and 2 are of the simple coast type, while ROI 3 comprises the complex coast type. Due to the vastness of the study are, some cloud pixels could be present in the images even when the sky is clear. Therefore, each ROI was defined as only a narrow vicinity of land and sea to minimize the number of cloud pixels over the study area and especially over the coastline. This vicinity covers enough area to apply segmentation properly and delineate the coastlines while reducing the challenges associated with cloud pixels.

Coastline monitoring requires time-series data due to the dynamic behavior of coastal regions, and the temporal interval for data

Table 1
Landsat-8T1 L2 images used to generate the Caspian Sea dataset.

Acquisition Date	Cloud Cover (%)	Path	Row	ROI
2013-06-03	1	165	034	ROI 1
2014-01-13	1.85			
2016-02-04	2.59			
2017-01-05	1.37			
2019-08-23	0.01			
2020-12-15	4.71			
2013-10-16	2.19	166	034	ROI 2
2014-08-16	0.6			
2015-04-29	0.49			
2016-08-21	0.02			
2017-08-08	0			
2018-03-20	1.65			
2013-08-29	0.26	166	032	ROI 3
2014-08-16	0.32			
2015-04-29	0.15			
2016-08-21	0			
2017-08-08	0			
2018-07-10	1.47			

acquisition is dependent on the context of the study (Beach et al., 2005). Acquiring and processing time-series satellite images is time-consuming and labor-intensive, making the data preparation step difficult. Google Earth Engine (GEE) offers a perfect solution for this problem by making it possible to access and process satellite images from multiple sources in a cloud computing platform. GEE also makes it possible to crop the necessary regions from the images and download the resulting patches for local processing, reducing the dataset volume dramatically.

In this study, we selected a Landsat-8 Tier-1 Level 2 image per year with less than 5% cloud cover from 2013 to 2020 within GEE. This interval provided enough data for model testing and coastline variability analysis purposes. Even though the cloud cover was set to a low value, i. e., 5%, while selecting full-frame images, there were no images with cloud-free coastlines in some years. Thus, there were no samples for some years in each ROI in the Caspian Sea dataset. Finally, ROI's shapefiles cropped the selected images, and the Blue-Green-Red-NIR bands were stacked and exported. There were six four-band images for each ROI with their corresponding ground truth images manually annotated. This process involved creating a land region shapefile by delineating the coastline over the RGB image in QGIS software and subsequently converting the resulting shapefile to a binary raster. A list of all the selected images in the Caspian Sea dataset is presented in Table 1.

2.2. Datasets preparation

2.2.1. China dataset

Ground truth images in the downloaded benchmark dataset were of RGB type with red and black regions representing land and sea areas. Utilizing this format would result in redundancy and incompatibility with the designed models. Therefore, the ground truth images have been converted into a single-band binary type with white and black regions representing land and sea areas. Using the Green and NIR bands (Landsat-8 bands 3 and 5 respectively) available in this dataset, NDWI images have been calculated for all the samples using Eq. (1). In total, three image sets have been selected from the dataset: RGB, NIR, and NDWI. Upon investigating the downloaded dataset, it was revealed that most of the samples included only one class, either sea or land. As data redundancy, these samples were excluded from the dataset. As a result, only 831 samples remained from the training and testing datasets. 78 of 831 remaining samples were selected for testing and the rest for training (Fig. 3). Like the Caspian Sea dataset (Section 2.1.2.), three coast types have been considered while selecting testing samples to evaluate trained models' performance better. These coast types include morphologically and geometrically simple and complex coasts, and small islands surrounded by water. More than half of the samples in the China testing dataset had complex coast types.

$$NDWI = \frac{Green - NIR}{Green + NIR} \tag{1}$$

Due to the exclusion of single-class samples, the data volume dwindled to about a quarter of its initial size. Thus, data augmentation techniques can synthetically increase the training dataset size in this situation. More training data can help prevent potential overfitting issues, leading to a better generalization (Yu et al., 2017). Data augmentation techniques include random rotation, flipping, noise injection, blurring, and rescaling (Shorten and Khoshgoftaar, 2019). The augmentation techniques that alter the pixel value or change the position of the edges could not be used in this study due to the importance of the coastline position. Therefore, the training dataset augmentation was achieved by applying every possible flip operation to the images and
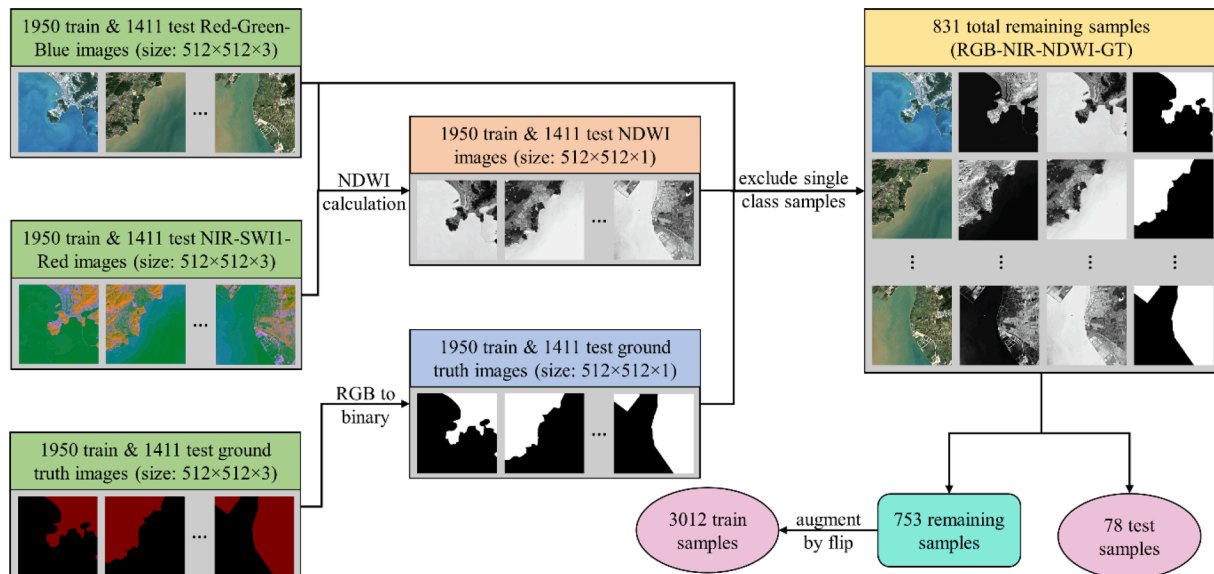

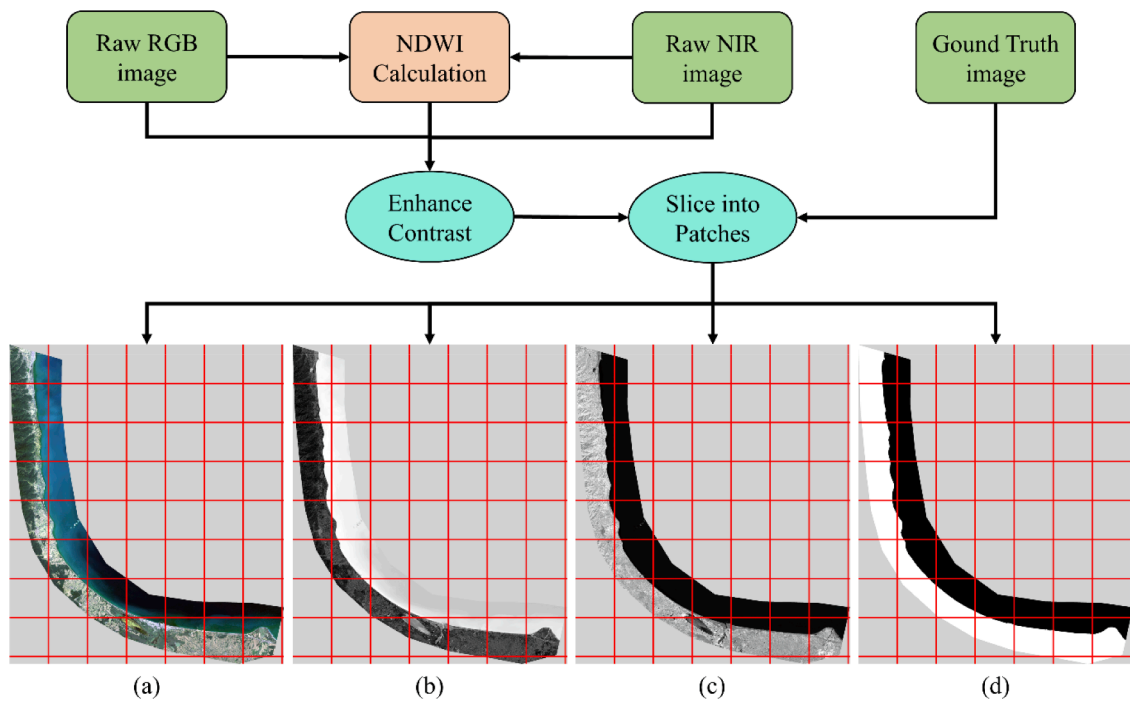**Fig. 3.** China dataset preparation steps.

Fig. 4. Caspian Sea dataset preprocessing steps. The depicted sample is associated with ROI 2 and the gray areas are No-Data. (a) RGB; (b) NDWI; (c) NIR; (d) GT.

their corresponding ground truth images. This process increased the training dataset size by a factor of four.

2.2.2. Caspian Sea dataset

The Caspian Sea testing dataset needed some preprocessing before being utilized. Firstly, NDWI images have been calculated using the raw images in the same manner as the China dataset. Most Landsat-8 images have low contrast due to the low reflectance of natural objects on the earth's surface. Thus, all the collected samples in the Caspian Sea dataset have been enhanced by the linear 1% stretch method. Images associated with ROIs 1, 2, and 3 have a size of 4196 × 3607, 3614 × 5196, and 2757 × 3052 pixels, respectively. These image sizes are too large to be

fed into the models directly, so they have been sliced into patches of 512 × 512 pixels in the final preprocessing step. After the prediction step, these patches would be attached to form the full segmentation maps of the same size as the original images. A flowchart of the preprocessing steps of the Caspian Sea dataset can be seen in Fig. 4.

3. Methodology

3.1. Sea-land segmentation models

U-Net is one of the most popular CNNs, first introduced by Ronneberger et al. (2015). It was initially designed for biomedical image
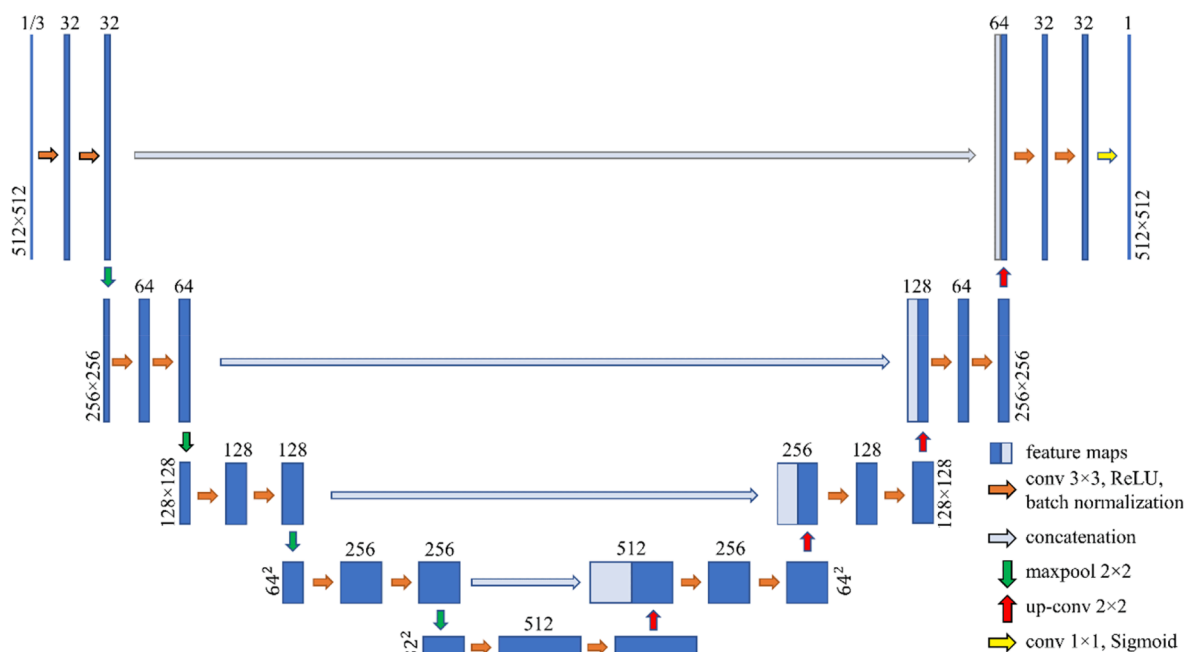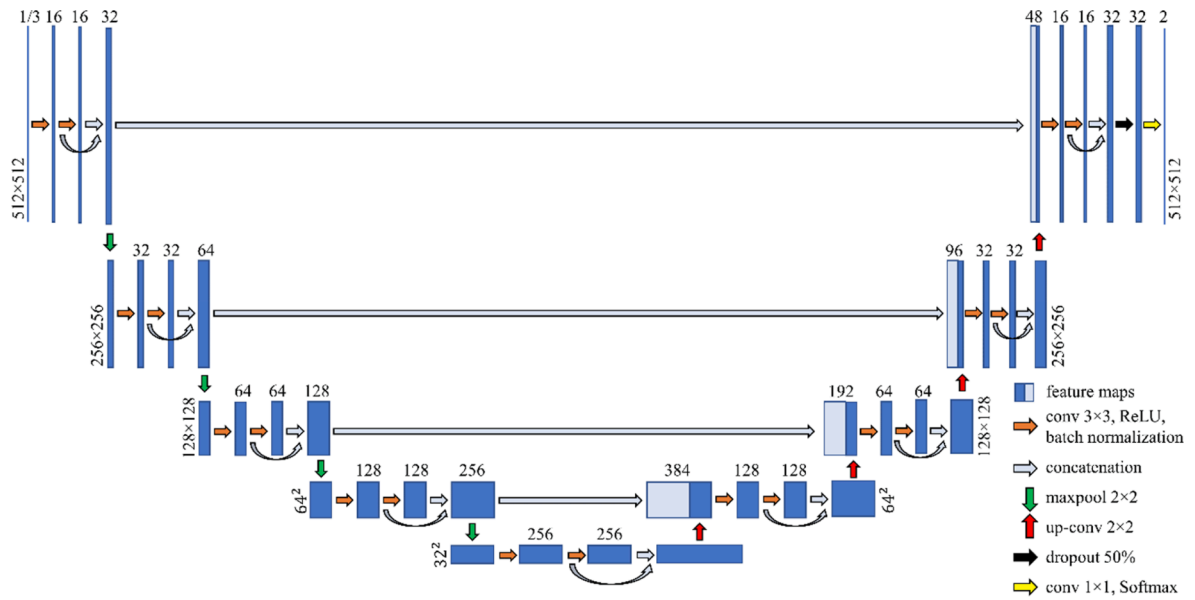

Fig. 5. Model 1 architecture.

Fig. 6. Model 3 architecture.

segmentation purposes. However, it has also been utilized in many other fields due to its adequate performance. The SUN architecture features four encoder and decoder blocks individually and a bottleneck block. Each block's activation functions are the rectified linear (ReLU). It utilizes 3 × 3 unpadded filters for convolution and 2 × 2 filters for both up-convolution and max-pooling. The number of convolution filters in the first layer is 64, doubling every subsequent encoder block until the bottleneck block. After that, the number of convolution filters halves in every decoder block. There is a 1 × 1 convolution operation at the final layer mapping the last feature vector to the specified classes. Three models are proposed based on the SUN architecture for the SLS task in this study.

3.1.1. Model 1

Model 1 is the SUN trained to set the basis for evaluating other models. The only addition to the SUN architecture in this model is the employment of a Batch Normalization (BN) layer (Fig. 5) before the

ReLU activation which has become an essential practice in DL (Luo et al., 2019). It has been proven that BN enhances generalization and convergence in the training process (Ioffe and Szegedy, 2015). This model takes either single-band NDWI or three-band RGB images as the input and the number of convolution filters in the first layer is 32. Activation functions in all the layers are the ReLU except in the final 1 × 1 convolutional layer, which is the Sigmoid. Finally, the Adam optimizer compiled the model with a learning rate of 0.001 and the binary cross-entropy loss function.

3.1.2. Model 2

Model 2 takes the architecture in Model 1 and tries to maximize its performance by modifying its elements. In this model, the NIR and RGB images are concatenated in an early fusion manner, i.e., before the encoding step, to form four-band images passing into the model as the input. In addition, a dropout layer has been added to the final decoder block to prevent potential overfitting issues. The significant enhance-
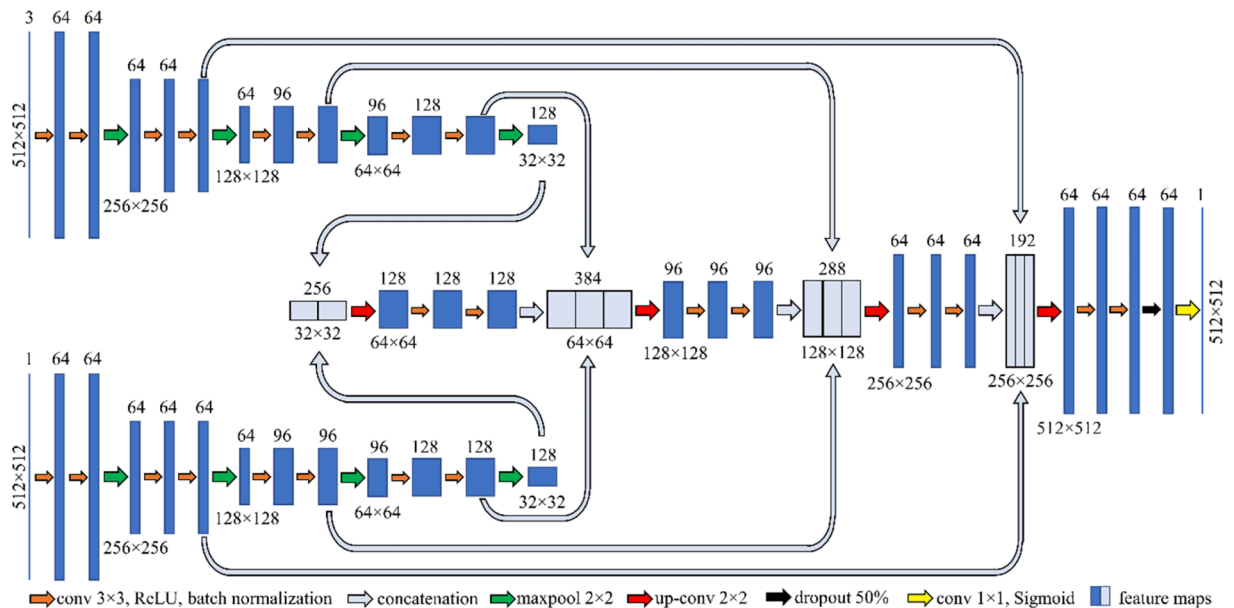

Fig. 7. Model 4 architecture.

ment in this model compared to Model 1 is a custom loss function based on the Jaccard coefficient. The problem with the binary cross-entropy loss function is that it could be saturated when the accuracies are high due to class imbalance. The cross-entropy loss value is usually very low for a given sample with a much smaller target area than the background. This problem can cause an unrealistic representation of the training data, resulting in inefficient loss value estimation. Assuming y and $\widehat{y}$ to be the $n \times m$ ground truth and prediction images, respectively, the Jaccard loss function for a given sample is defined using Eq. (2).

$$Loss_{IoU} = -\frac{\sum_{i=1}^{n}\sum_{j=1}^{m} y_{i,j}.\widehat{y}_{i,j}}{\sum_{i=1}^{n}\sum_{j=1}^{m}\left(y_{i,j} + \widehat{y}_{i,j} - y_{i,j}.\widehat{y}_{i,j}\right)} \qquad (2)$$

3.1.3. Model 3

Model 3 is designed to reduce the training time while improving the performance compared to the SUN. For this purpose, the model depth has been decreased by setting the number of convolutional filters in the first layer to 16, which dramatically reduces the number of trainable parameters. Two workarounds have been suggested to compensate for the performance loss caused by this decision. Firstly, a concatenation operation is introduced within each convolutional block similar to the Residual blocks in the ResNet (He et al., 2016), as can be seen in Fig. 6. This operation should help further improve the feature extraction process within each block. Secondly, a custom weighted binary cross-entropy loss function has been incorporated in this model. It has been proven that dedicating a higher probability to the foreground targets in the cross-entropy loss function makes learning imbalanced samples more efficient (Aurelio et al., 2019). Like Model 1, Model 3 is also designed to take either NDWI or RGB images as the input. A dropout layer is present in this model, and the activation function for the final convolution layer is the Softmax. Other than the loss function, this model is compiled similarly to the previous ones.

3.1.4. Model 4

Model 4′s architecture is inspired by the FuseNet model proposed by Yang et al. (2020a). This model has a relatively lower width than the previous models, hence the reduced number of trainable parameters. It is a dual encoder network with different inputs for each path that concatenates the inputs in a late fusion manner, i.e., after the encoding step. The idea is to extract features from different inputs separately and combine them into a more informative feature set in the decoding process. As shown in Fig. 7, skip features from the encoder blocks are concatenated with the up-convolution outputs in the decoder path. In this model, the RGB and NIR images were passed into the first and second encoder paths, respectively, and were fused after encoding. Model 4 is compiled the same way as the other models, and it also features the custom Jaccard loss function introduced in Model 2.

3.2. Training strategies and comparison with other segmentation models

Two training scenarios have been considered for each model defined in Section 3.1. for the ablation study on input data types. Conversely, the ablation study analyzed the loss functions when the input data type was unchangeable. Model 1 was firstly trained on RGB images to lay the foundation for comparing different training strategies. This scenario is referred to as the SUN in this paper. Model 1 was then trained on the single-band NDWI images. Model 2 comprised two loss functions (i.e., Cross-entropy and Jaccard) used in two individual scenarios. Model 3 was first trained on RGB then NDWI images, as in Model 1. Moreover, Model 4 consisted of the same training scenarios as Model 2.

Two other models have also been implemented for a comprehensive assessment of our methods, including FC-DenseNet (Jegou et al., 2017) and DeepLabV3+ (Chen et al., 2018). These models are among the most commonly used state-of-the-art semantic segmentation models that have also been successfully employed in SLS (Yang et al., 2020b). Specifications of different training strategies can be seen in Table 2.

Table 2
Specifications of different training strategies. (Sc. stands for Scenario).

Models	Input data	Number of trainable parameters	Train time per epoch (seconds)	Loss Function
FC-DenseNet	RGB	3,457,872	1,482	Cross-entropy
DeepLabV3+	RGB	11,819,874	793	Cross-entropy
Model 1 Sc. 1	RGB	7,765,985	236	Cross-entropy
Model 1 Sc. 2	NDWI	7,765,409	229	
Model 2 Sc. 1	RGB + NIR	7,766,273	240	Cross-entropy
Model 2 Sc. 2				Jaccard
Model 3 Sc. 1	RGB	3,140,898	162	Cross-entropy
Model 3 Sc. 2	NDWI	3,140,610	161	
Model 4 Sc. 1	RGB, NIR	2,032,481	421	Cross-entropy
Model 4 Sc. 2				Jaccard

3.3. Automatic coastline extraction

Edge detection methods can automatically extract the coastline from the segmentation outputs. However, directly applying these techniques would cause two significant problems. Firstly, the segmentation errors can create defects in the edge image unless the segmentation image has an IoU score of 100%, which hardly ever happens in any segmentation problem. Secondly, the ROI boundary also gets extracted as the edge which is not of interest and must be filtered out. This study has developed a post-processing pipeline to overcome these issues, as presented in Fig. 8. The post-processing takes the segmentation image, a structuring element designed according to the 4-neighbors pixel connectivity, and the binary ROI mask as the input and extracts the coastline in 8 steps.

The processing starts with step 1, in which the initial edges are detected in the segmentation image. The edge detection process is carried out by applying a morphological erosion operation to the segmentation image using the designed structuring element and subtracting the erosion image from the segmentation image. The resulting edge image contains an enclosed region consisting of the coastline and the lower ROI boundary and some redundant edges caused by the defects in the segmentation image. It should be noted that the detected lower ROI boundary is within the binary ROI mask and cannot be filtered out. In step 2, the redundant edges are excluded by pixel connectivity analysis. The idea is to classify white pixels based on their connectivity of 8-neighbors. Therefore, all the connected land pixels are identified as an individual class. After the classification, the class with the most pixels corresponds to the enclosed region, while the classes corresponding to the redundant edges have a much lower pixel count. This information can be used to effectively exclude all the redundant edges that are not connected to the lower ROI boundary or the coastline. It is noteworthy that the 8-neighbors connectivity is considered for this classification to prevent the removal of the sea-land interface pixels that are only diagonally connected to the other land pixels.

Step 3 starts by filling the black region within the white enclosed area by the flood-fill algorithm. The structuring element used for erosion ensures that the black region within the enclosed area has 4-neighbors connectivity. Thus, all the extracted edge pixels, including those that were only diagonally connected, are attached to the filled region, resulting in the 4-neighbors connectivity between all the land pixels. Due to the extra care taken during the image collection process in the Caspian Sea dataset, the sea-land interface in the images was cloud-free. However, some cloud pixels near the ROI boundary could not be avoided. Therefore, the extracted coastlines could not have had any imperfections, but some imperfections could be present in the lower ROI boundary. In step 4, these imperfections are made separable by using the binary ROI mask to fill the No-Data region in the output image of the previous step with white pixels. If imperfection is present in the lower
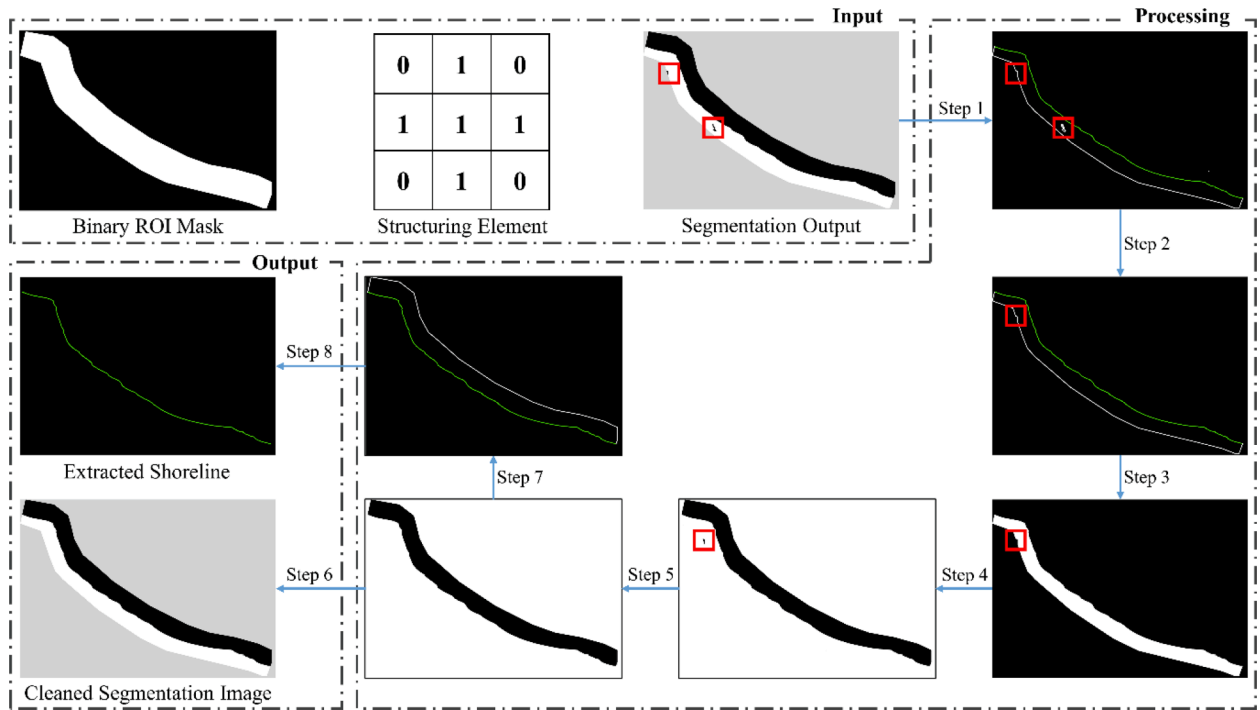
Fig. 8. Automatic coastline extraction flowchart. The green line and the red squares depict the coastline and segmentation defects, respectively, and the gray areas are No-Data. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

ROI boundary, it is revealed as a black pixel cluster. These clusters are eliminated in step 5 by applying the pixel connectivity classification on the black pixels with the 4-neighbors connectivity method. Keeping the class with the highest pixel count and excluding any other classes could efficiently eliminate the imperfections connected to the lower ROI boundary.

At this point, all the imperfections in the segmentation image are removed, and the black and white regions are perfectly separable, each having the 4-neighbors connectivity. The cleaned segmentation image is easily generated in step 6 by applying the binary ROI mask to the output from step 5. For extracting the coastline, firstly, the edges in the output from step 5 are detected in step 7. The edge detection method is the same as in step 1. The resulting edge image contains the coastline and the upper ROI boundary. Finally, the coastline is isolated in step 8 by excluding the upper ROI boundary. Since the detected upper ROI boundary is out of the binary ROI mask, it can be filtered out easily. It should be noted that the detected coastline pixels belong to the land class. The output binary coastline raster can be converted to the coastline shapefile with raster to polyline toolboxes.

3.4. Evaluation metrics

Outputs of the models with the Sigmoid activation function were single-band images representing the probability of each pixel belonging to the land class. Thus, the probabilities were thresholded by a value of 0.5 to generate the final binary segmentation images. On the other hand, outputs of the models with the Softmax activation function had two bands corresponding to the land and sea classes. In this case, the final segmentation images were generated by assigning the class with the maximum prediction value to each pixel. A confusion matrix can be calculated by comparing the resulting binary segmentation images to the ground truth. Since all the images in the Caspian Sea dataset had No-Data areas, it was assured that only the ROI region in the segmentation image was compared to its corresponding region in the ground truth.

From the confusion matrix, several measures are extracted, including True Positive (TP), True Negative (TN), False Positive (FP), and False

Negative (FN) values. These measures define the following metrics for assessing the models' performance: Accuracy, Sensitivity, Precision, F1-score, and IoU. Among these five metrics, IoU is the most strict and, it usually has a lower value than the others. It considers the overlap between the prediction and ground truth images and penalizes both the false positives and the false negatives. These five evaluation metrics are defined in Eq. (3) to (7).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{3}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{4}$$

$$Precision = \frac{TP}{TP + FP} \tag{5}$$

$$F1 - Score = \frac{2.Precision.Sensitivity}{Precision + Sensitivity} \tag{6}$$

$$IoU = \frac{TP}{TP + FP + FN} \tag{7}$$

3.5. Implementation

Due to the hardware limitations, training DL models can be challenging on local machines (Ball et al., 2017). On the other hand, programming environment maintenance for DL can be equally challenging sometimes. The development of free cloud computing services like Google Colaboratory (Carneiro et al., 2018) is auspicious due to the availability of free accelerators like TPUs (Tensor Processing Unit) and GPUs (Graphical Processing Unit). Combined with the Google Drive service, it is ideal for training medium to relatively large-sized models.

The preprocessed data were uploaded to Google Drive, then accessed within the Google Colaboratory notebook. All the developed models in Section 3.1 were designed with the Keras and Tensorflow API within the Python environment. The training was carried out for 150 epochs in all the models with a Tesla T4 GPU selected as the accelerator. 10% of the training

Table 3
China dataset quantitative results (maximum values are bold).

Models	Mean Accuracy (%)	Mean Sensitivity (%)	Mean Precision (%)	Mean F1-Score (%)	Mean IoU (%)
FC-DenseNet	97.07	92.03	95.97	92.97	88.65
DeepLabV3+	99.23	94.72	**98.09**	95.66	93.15
Model 1 Sc. 1	99.35	94.53	96.74	94.90	92.77
Model 1 Sc. 2	99.10	95.15	97.85	96.13	93.67
Model 2 Sc. 1	99.34	95.71	97.85	96.36	93.84
Model 2 Sc. 2	**99.42**	**96.16**	96.80	96.40	**94.45**
Model 3 Sc. 1	99.28	96.06	96.49	96.20	94.08
Model 3 Sc. 2	99.01	94.94	97.85	95.76	93.18
Model 4 Sc. 1	99.09	96.12	97.33	**96.47**	93.75
Model 4 Sc. 2	99.35	95.84	97.57	95.99	93.95

samples were selected just for validation purposes to monitor overfitting issues while training. It should be noted that only the China dataset was used for training. The data and the source codes developed in this study will be openly available upon publication at https://github.com/mohammad-aghdaminia/sea-land-segmentation-coastline-extraction.

4. Results and discussion

In this study, different training strategies have been evaluated based on their IoU scores. This metric offers a more realistic representation of the models' performance and usually has a lower value than the other metrics. The two testing datasets in this study were assessed separately in the following two sub-sections.

4.1. China dataset

The trained models were first evaluated on the China dataset. Table 3 represents the quantitative results of the China dataset. Model 2 Sc. 2 achieved the best performance among all the models, offering the highest IoU, Accuracy, and Sensitivity scores of 94.45%, 99.42%, and 96.16%, respectively, while FC-DenseNet yielded the lowest scores in all the metrics. Moreover, the highest F1-Score and Precision scores (96.47% and 98.09%, respectively) were associated with DeepLabV3+ and Model 4 Sc. 1. On the other hand, Model 1 performed better with NDWI images (Sc. 2) than with RGB images (Sc. 1), with the difference in the IoU score being 0.9%. Unlike Model 1, Model 3 performed the best on RGB images (Sc. 1), offering the second-highest IoU score.
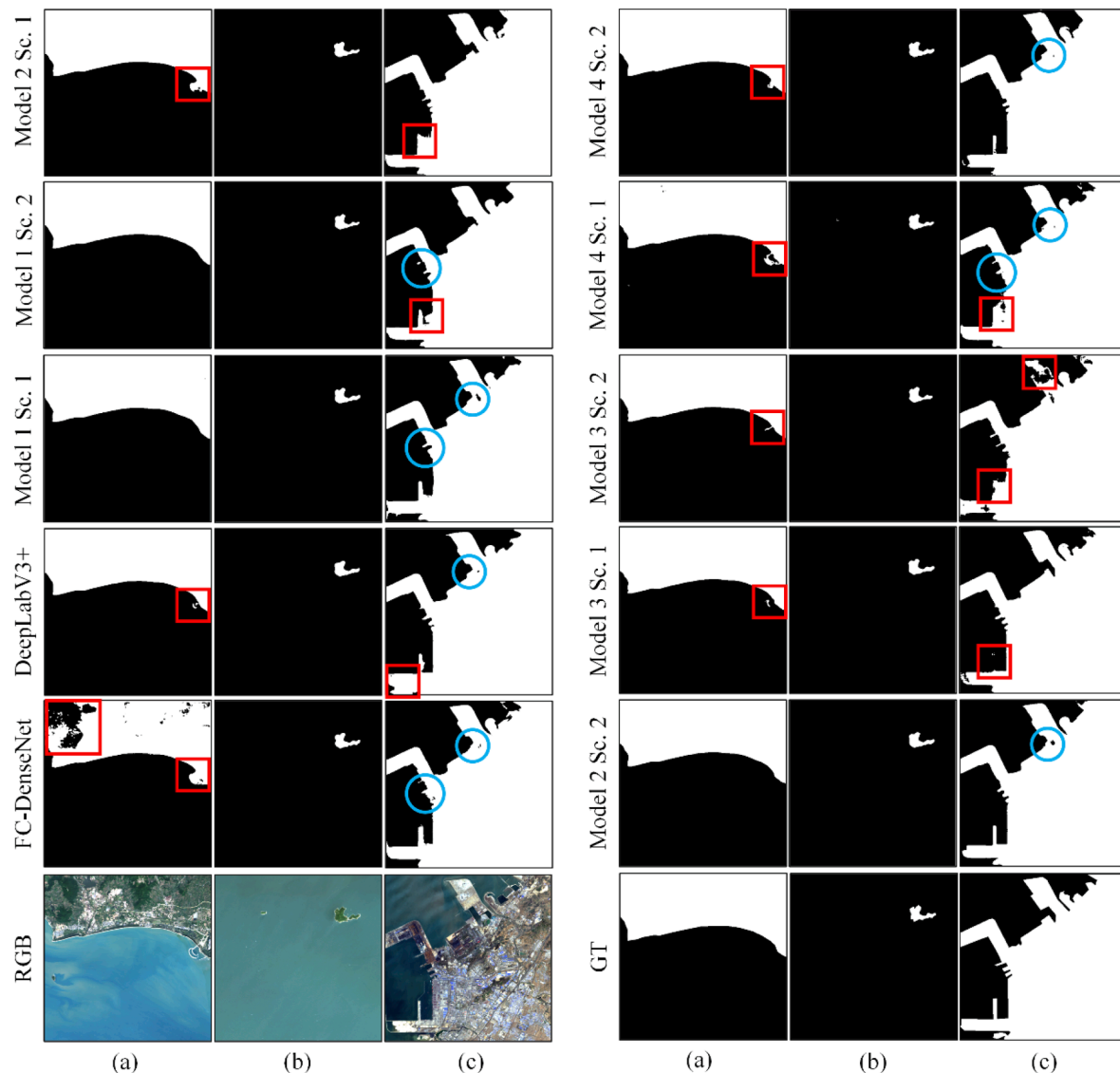

Fig. 9. Visual results of the China dataset. Blue circles and red squares represent detail loss and artifact, respectively. (a) simple coast; (b) small island surrounded by water; (c) complex coast. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 4
Caspian Sea dataset quantitative results. (maximum values are bold).

Models	Mean Accuracy (%)	Mean Sensitivity (%)	Mean Precision (%)	Mean F1-score (%)	Mean IoU (%)
FC-DenseNet	92.42	85.67	98.99	91.44	84.92
DeepLabV3+	80.44	70.75	87.61	77.61	64.72
Model 1 Sc. 1	94.62	90.94	98.62	94.37	89.92
Model 1 Sc. 2	98.79	97.80	99.73	98.75	97.54
Model 2 Sc. 1	99.13	98.53	99.76	99.14	98.29
Model 2 Sc. 2	**99.43**	**99.05**	**99.82**	**99.43**	**98.87**
Model 3 Sc. 1	95.94	92.82	99.10	95.81	92.09
Model 3 Sc. 2	98.82	98.01	99.59	98.79	97.61
Model 4 Sc. 1	97.75	96.08	99.56	97.76	95.69
Model 4 Sc. 2	99.00	98.24	99.76	98.99	98.01

Furthermore, Sc. 2 in Model 4, which used the Jaccard loss function, provided better results than Sc. 1. All the models had better performance than Model 1 with RGB input (Sc. 1), defined as the base model for comparisons. These results indicate that all of the proposed modifications to the SUN architecture could increase the performance. The most significant improvement to the IoU score was as much as 1.68%. Moreover, our modified models outperformed FC-DenseNet and DeepLabV3+ networks.

Fig. 9 presents the visual results for a sample from each coast type defined in Section 2.2.1. Most models, except FC-DenseNet, usually performed well on the samples with simple coasts. The predictions were very close to the ground truth images in these samples. However, some models (e.g., DeepLabV3+ and Model 4 Sc. 1) introduced slight artifacts to the predictions that were negligible in most cases. Moreover, samples with small islands also lead to satisfactory predictions. Although the target area in these samples was much smaller than the background area, all the models predicted these samples correctly. The major challenge was with the samples containing complex coasts. Artifacts were the most dominant in these samples, and there was also detail loss that was not a big issue with the previous coast types. Detail loss mainly occurred in small areas with higher geometric complexity. Model 2 Sc. 2 offered the least artifacts and detail loss in this coast type.

4.2. Caspian Sea dataset

The next step in the evaluation process was to test the trained models on the Capsian Sea dataset. Table 4 represents the quantitative results of the Caspian Sea dataset. Similar to the China dataset, Model 2 Sc. 2 achieved the best results, yielding the highest values in all the metrics (IoU = 98.87%, Accuracy = 99.43%). In addition, Model 2 Sc. 1
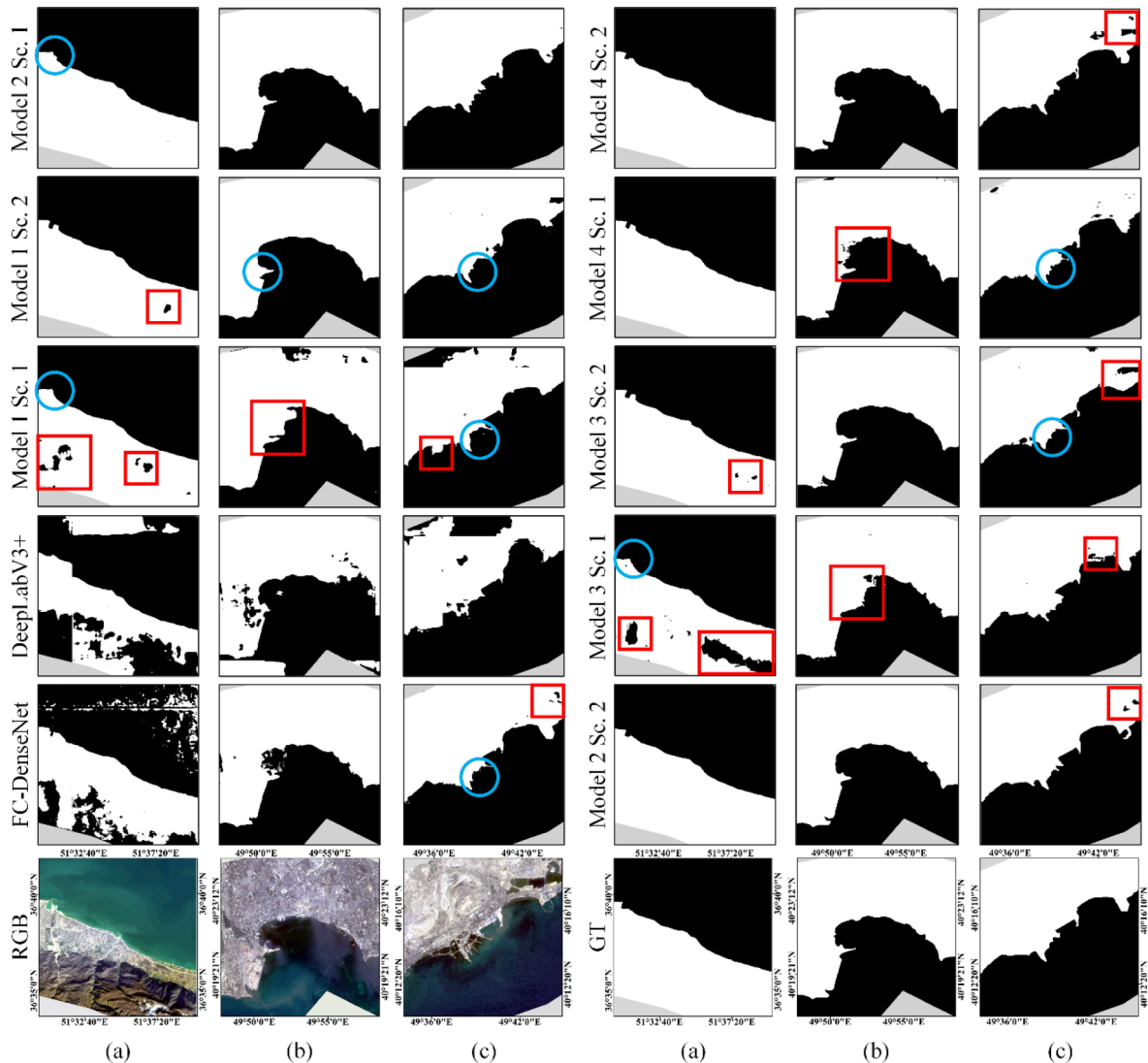

Fig. 10. Visual results of the Caspian Sea dataset. Blue circles and red squares represent detail loss and artifact, respectively, and the gray areas are No-Data. (a) simple coast; (b-c) complex coast. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

performed on par with Model 2 Sc. 2, offering the second-best scores in all the statistical indices, while DeepLabv3+ provided the lowest performance (IoU = 64.72%, Accuracy = 80.44%). On the other hand, Models 1 and 3 performed better with NDWI images (Sc. 2), with the latter having a slightly higher IoU score. Moreover, Model 4 had better results in Sc. 2 (Jaccard loss), increasing the IoU score up to 2.32% compared to Sc. 1. Our models enhanced the performance compared to the SUN, with the most significant improvement to the IoU score being as high as 8.95%. FC-DenseNet and DeepLabV3+ could not generate comparable results to our models in this dataset.

Fig. 10 depicts the visual results of a sample with the simple (Sample a) and two samples with the complex (Samples b and c) coast types in the Caspian Sea dataset. The results indicated that the FC-DenseNet and DeepLabV3+ models could not delineate the sea and land classes with a high level of accuracy and produced the worst segmentation outputs. As predicted from the quantitative results (Table 4), the SUN (i.e., Model 1 Sc. 1) resulted in the most misclassifications compared to our modified models in all the coast types. Model 2 Sc. 1 and 2 offered the most accurate sea-land maps based on visual interpretation, with the latter having the cleanest outputs. Despite the few defects in the segmentation images of this model, none were near the coastline region, which is of great importance in the automatic coastline extraction process.

4.3. Comparison analysis

Six different DL models were evaluated in a previously published paper by Yang et al. (2020b), from which the China dataset in our study is derived. U-Net, one of these models, had achieved an IoU score of 93.11% in the best scenario. The best result in their study belonged to the DeepLabV3+ and FC-DenseNet models with IoU scores of 93.36% and 92.85%, respectively. Using the same dataset as in their study, we implemented these three models, the outputs of which were in agreement with their study. Furthermore, we developed seven other models that were modified forms of the SUN. Six out of these seven models in our study outperformed their best result (i.e., DeepLabV3+), with our highest IoU score being as much as 94.45% which belonged to Model 2 Sc. 2. These results proved that not only the enhancements to the SUN in our study were effective, but they also managed to outperform other previously deployed models as well.

Our study also featured a novel testing dataset from the southern Caspian Sea coastal region. The quantitative segmentation results indicated the poor generalizability of DeepLabV3+ and FC-DenseNet when facing unseen images from other test sites. Similarly, the SUN could not generalize well either, however, it was substantially better than the former two models in this regard. On the other hand, DeepLabV3+ and FC-DenseNet were significantly more time-consuming to train than the SUN (Table 2). Therefore, the SUN was selected as the base model for the enhancements developed in this study.

The evaluations on Models 1 and 3 in the two datasets revealed that utilizing NDWI images (Sc. 2) instead of RGB (Sc. 1) achieved better results in three out of four testing instances. The temporal efficiency introduced by NDWI images in the training phase was negligible. However, the main advantage of using NDWI instead of RGB images, besides the improved performance of the models, was reducing the training dataset volume by a factor of three, which is immensely valuable when dealing with hardware limitations. Models 2 and 4 were tested with the Cross-entropy (Sc. 1) and the Jaccard (Sc. 2) loss functions. In all four testing instances, the Jaccard loss function proved to be more efficient, introducing remarkable enhancement.

4.4. Coastline extraction

Using the prediction images of the 18 samples in the Caspian Sea dataset, it is possible to evaluate the coastal area and coastline changes in a time-series manner. For this purpose, the best model should be determined for the subsequent automatic coastline extraction using the

Table 5
Caspian Sea dataset quantitative results with a buffer size of 5 pixels. (maximum values are bold).

Models	Mean Accuracy (%)	Mean Sensitivity (%)	Mean Precision (%)	Mean F1-score (%)	Mean IoU (%)
FC-DenseNet	76.49	80.32	78.71	75.31	65.83
DeepLabV3+	73.48	68.01	71.44	66.36	58.56
Model 1 Sc. 1	82.10	78.75	87.99	79.87	71.35
Model 1 Sc. 2	77.87	64.33	90.31	71.24	61.22
Model 2 Sc. 1	86.84	83.56	93.35	86.01	78.19
Model 2 Sc. 2	**90.79**	**89.99**	**93.85**	**90.69**	**84.63**
Model 3 Sc. 1	83.66	88.05	83.67	83.67	75.89
Model 3 Sc. 2	80.42	70.37	90.45	75.58	66.27
Model 4 Sc. 1	85.01	84.75	90.02	84.76	76.49
Model 4 Sc. 2	89.22	88.27	92.72	88.95	82.38

post-processing step defined in Section 3.3. It should be noted that the sea-land interface constitutes a very small proportion of the pixels in the image; thus, very high evaluation metric values could be yielded even when there are disagreements between the actual sea-land interface and that generated by the models. Hence, the performance metrics were applied to the pixels belonging to the actual coastline in isolation rather than the whole image, considering a buffer window with a size of 5 pixels.

As is clear from Table 5, the closest segmentation outputs in the coastline buffer area to the ground truth images belonged to Model 2 Sc. 2. This model achieved the highest scores in all the evaluation metrics; therefore, its outputs were selected to be used in the coastline extraction process. The post-processing outputs were the cleaned SLS images and the coastline position polylines. The extracted coastlines with a zoomed area for each ROI can be seen in Fig. 11. For example, the zoomed image in ROI 2 (Fig. 11c and d) is a manufactured object with more change than the other two zoomed images, than the other two zoomed images which are natural features. It should be noted that there was no data for ROIs 2 and 3 after 2018.

It is possible to analyze the Caspian Sea coast area change quantitatively by multiplying the number of the land pixels to a single Landsat-8 ground pixel area which is 30×30 m squared. The calculated coastal area in each sample of ROIs 1 to 3 is subtracted from the corresponding coastal area in the 2013 samples to evaluate the land area change. The coastal area in each ROI has increased with respect to 2013, indicating land gain in all the years (Fig. 12). The land gain in ROIs 1 to 3 in the last-time samples (2020, 2018, and 2018, respectively) is 9.73, 9.87, and 1.98 km^2.

5. Conclusion

This study aimed to improve the SLS performance of the SUN by modifying its architecture and utilizing different input image types. Three different models have been proposed and compared to the SUN, which enhanced the performance. These models also outperformed two other state-of-the-art networks including FC-DenseNet and DeepLabV3+. Moreover, a post-processing pipeline has been developed for the automatic coastline extraction on the segmentation outputs. This procedure is based on edge detection by morphological erosion operation and subsequent defect elimination by pixel connectivity classification. The key findings of this study can be described as follows:

- Utilizing single-band NDWI images instead of three-band RGB images as the input often yielded better results while significantly reducing the training dataset volume by a factor of 3. However, the reduction in training time was negligible.
- The fusion of the NIR and RGB images improved the performance with the early fusion method outperforming the late fusion.
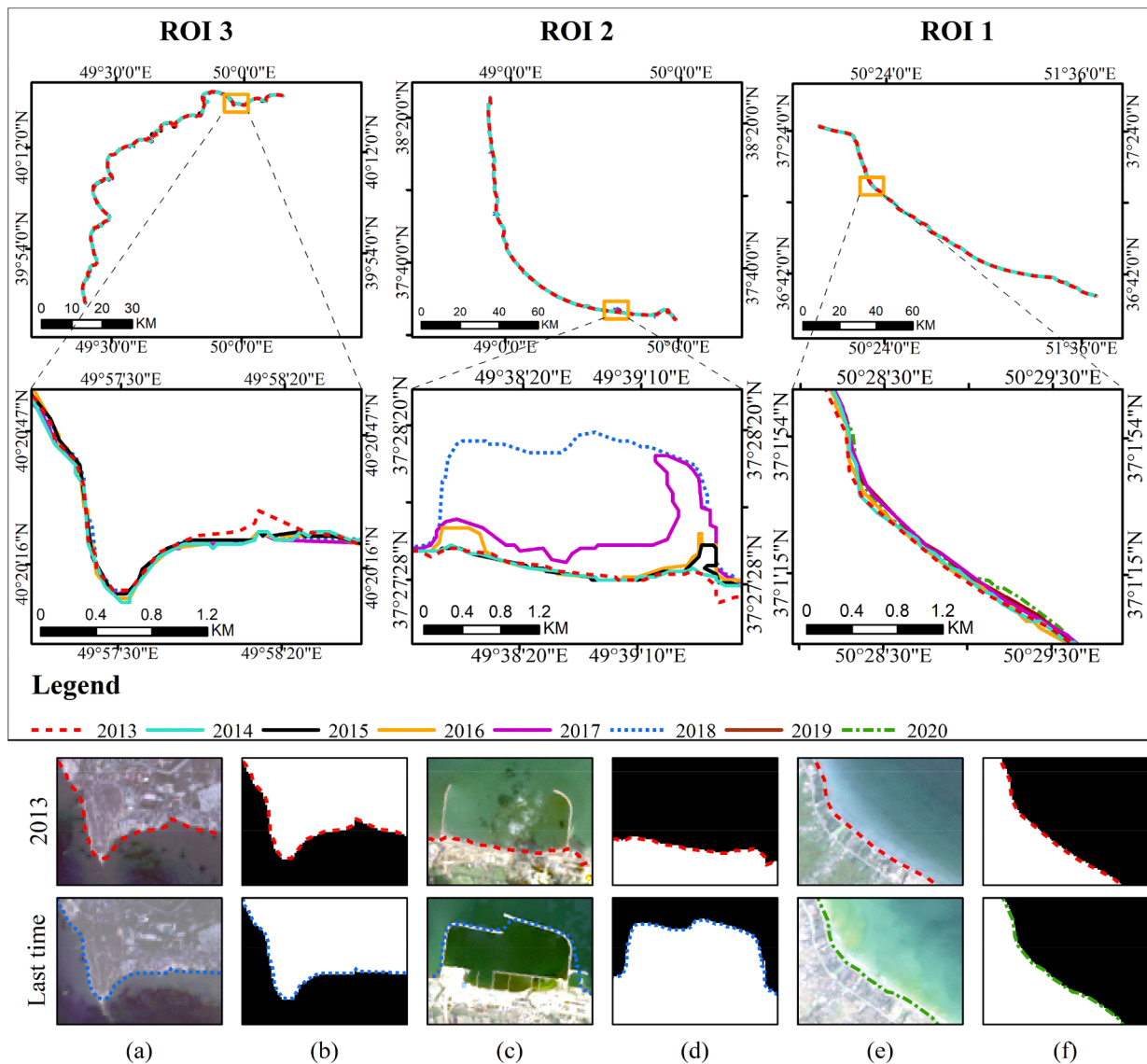
Fig. 11. The upper panel depicts the extracted coastline time-series of ROIs 1 to 3. The lower panel represents the RGB images and segmentation outputs of the zoomed area in a-b) ROI 1, c-d) ROI 2, and e-f) ROI 3.
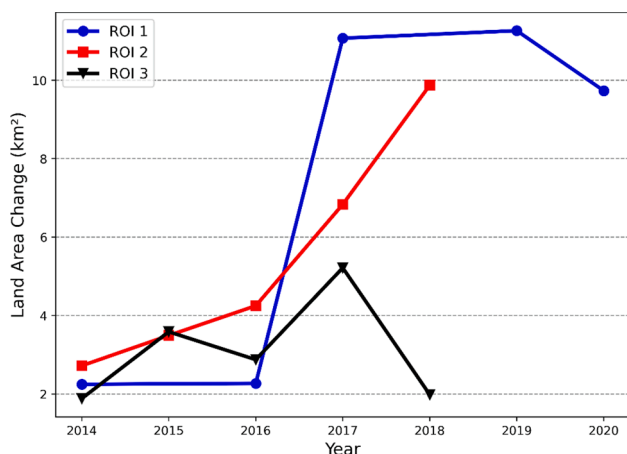

Fig. 12. Land area change in ROI 1 (blue circles), ROI 2 (red squares), and ROI 3 (black triangles) in each year compared to 2013. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

- The Jaccard loss function outperformed the cross-entropy loss function.
- The proposed automatic coastline extraction pipeline could efficiently delineate the coastlines with high accuracy while neglecting the defects.

Our study provides a fully automatic framework for the SLS of Landsat-8 coastal images and subsequent coastline extraction, although some limitations should be noted. The extracted coastlines from the Landsat-8 images have a resolution of 30 m. Therefore they might not be suitable for applications where finer resolutions are required. It was impossible to tackle this problem in our study due to the lack of a large open-source SLS benchmark dataset with a finer ground resolution, and it should be considered in relative studies. Future studies in this field could benefit from the land surface temperature (LST) and sea surface temperature (SST) products for better discrimination between sea and land bodies. Moreover, radar data could also offer helpful information due to the unique backscattering pattern of open water bodies.

CRediT authorship contribution statement

Mohammad Aghdami-Nia: Conceptualization, Methodology, Investigation, Software, Visualization, Writing – original draft, Writing – review & editing. **Reza Shah-Hosseini:** Conceptualization, Investigation, Project administration, Supervision, Writing – original draft. **Amirhossein Rostami:** Project administration, Visualization, Writing – review & editing. **Saeid Homayouni:** Investigation, Software, Visualization, 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.

References

Ansari, M., Homayouni, S., Safari, A., Niazmardi, S., 2021. A New Convolutional Kernel Classifier for Hyperspectral Image Classification. IEEE J Sel. Top. Appl. Earth Obs. Remote Sens. 14, 11240–11256. https://doi.org/10.1109/JSTARS.2021.3123087.

Aurelio, Y.S., de Almeida, G.M., de Castro, C.L., Braga, A.P., 2019. Learning from Imbalanced Data Sets with Weighted Cross-Entropy Function. Neural Process. Lett. 50 (2), 1937–1949. https://doi.org/10.1007/s11063-018-09977-1.

Ball, J.E., Anderson, D.T., Chan, C.S., 2017. Comprehensive survey of deep learning in remote sensing: theories, tools, and challenges for the community. J. Appl. Remote Sens. 11, 1. https://doi.org/10.1117/1.jrs.11.042609.

Beach, W.P., Boak, E.H., Turner, I.L., Street, K., Vale, M., 2005. Shoreline Definition and Detection : A Review 688–703. https://doi.org/10.2112/03-0071.1.

Carneiro, T., Medeiros Da Nobrega, R.V., Nepomuceno, T., Bian, G.-B., De Albuquerque, V.H.C., Filho, P.P.R., 2018. Performance Analysis of Google Colaboratory as a Tool for Accelerating Deep Learning Applications. IEEE Access 6, 61677–61685. https://doi.org/10.1109/ACCESS.2018.2874767.

Chen, L.C., Zhu, Y., Papandreou, G., Schroff, F., Adam, H., 2018. Encoder-decoder with atrous separable convolution for semantic image segmentation. Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics) 11211 LNCS, 833–851. https://doi.org/10.1007/978-3-030-01234-2_49.

Chen, W.-W., Chang, H.-K., 2009. Estimation of shoreline position and change from satellite images considering tidal variation. Estuar. Coast. Shelf Sci. 84 (1), 54–60. https://doi.org/10.1016/j.ecss.2009.06.002.

Cheng, D., Meng, G., Xiang, S., Pan, C., 2017. FusionNet: Edge Aware Deep Convolutional Networks for Semantic Segmentation of Remote Sensing Harbor Images. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 10 (12), 5769–5783. https://doi.org/10.1109/JSTARS.2017.2747599.

Cui, B., Jing, W., Huang, L., Li, Z., Lu, Y., 2021. SANet: A Sea-Land Segmentation Network Via Adaptive Multiscale Feature Learning. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 14, 116–126. https://doi.org/10.1109/JSTARS.2020.3040176.

Elkhateeb, E., Soliman, H., Atwan, A., Elmogy, M., Kwak, K.S., Mekky, N., 2021. A Novel Coarse-To-Fine Sea-Land Segmentation Technique Based on Superpixel Fuzzy C-Means Clustering and Modified Chan-Vese Model. IEEE Access 9, 53902–53919. https://doi.org/10.1109/ACCESS.2021.3065246.

Feyisa, G.L., Meilby, H., Fensholt, R., Proud, S.R., 2014. Remote Sensing of Environment Automated Water Extraction Index : A new technique for surface water mapping using Landsat imagery. Remote Sens. Environ. 140, 23–35. https://doi.org/10.1016/j.rse.2013.08.029.

Firoozfar, A., Bromhead, E.N., Dykes, A.P., Lashteh Neshaei, M.A., 2012. Southern Caspian Sea Coasts Morphology Sediment Characteristics and Sea Level Change. Proc. Annu. Int. Conf. Soils, Sediments, Water Energy 17, 29.

Giardino, C., Bresciani, M., Villa, P., Martinelli, A., 2010. Application of Remote Sensing in Water Resource Management: The Case Study of Lake Trasimeno. Italy. Water Resour. Manag. 24 (14), 3885–3899. https://doi.org/10.1007/s11269-010-9639-3.

Hamylton, S.M., Prosper, J., 2012. Development of a spatial data infrastructure for coastal management in the Amirante Islands, Seychelles. Int. J. Appl. Earth Obs. Geoinf. 19, 24–30. https://doi.org/10.1016/j.jag.2012.04.004.

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. 2016-Decem, 770–778. https://doi.org/10.1109/CVPR.2016.90.

Huihui, X., Qizhi, X., Lei, H., 2016. A sea-land segmentation algorithm based on Gray Smoothness Ratio, in: 4th International Workshop on Earth Observation and Remote Sensing Applications, EORSA 2016 - Proceedings. pp. 117–121. https://doi.org/10.1109/EORSA.2016.7552778.

Ioffe, S., Szegedy, C., 2015. Batch normalization: Accelerating deep network training by reducing internal covariate shift. 32nd Int. Conf. Mach. Learn. ICML 2015 1, 448–456.

Jegou, S., Drozdzal, M., Vazquez, D., Romero, A., Bengio, Y., 2017. The One Hundred Layers Tiramisu: Fully Convolutional DenseNets for Semantic Segmentation. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Work. 2017-July, 1175–1183. https://doi.org/10.1109/CVPRW.2017.156.

Kelly, J.T., Gontz, A.M., 2018. Using GPS-surveyed intertidal zones to determine the validity of shorelines automatically mapped by Landsat water indices. Int. J. Appl. Earth Obs. Geoinf. 65, 92–104. https://doi.org/10.1016/j.jag.2017.10.007.

Khoshboresh-Masouleh, M., Alidoost, F., Arefi, H., 2020. Multiscale building segmentation based on deep learning for remote sensing RGB images from different sensors. https://doi.org/10.1117/1.JRS.14.034503 14, 034503. https://doi.org/10.1117/1.JRS.14.034503.

Lei, S., Zou, Z., Liu, D., Xia, Z., Shi, Z., 2018. Sea-Land Segmentation for Infrared Remote Sensing Images based on Superpixels and Multi-scale Features $. Infrared Phys. Technol. 91, 12–17. https://doi.org/10.1016/j.infrared.2018.03.012.

Li, H., Zech, J., Ludwig, C., Fendrich, S., Shapiro, A., Schultz, M., Zipf, A., 2021a. Automatic mapping of national surface water with OpenStreetMap and Sentinel-2 MSI data using deep learning. Int. J. Appl. Earth Obs. Geoinf. 104, 102571. https://doi.org/10.1016/j.jag.2021.102571.

Li, K., Wang, J., Yao, J., 2021b. Effectiveness of machine learning methods for water segmentation with ROI as the label: A case study of the Tuul River in Mongolia. Int. J. Appl. Earth Obs. Geoinf. 103, 102497. https://doi.org/10.1016/j.jag.2021.102497.

Li, R., Liu, W., Yang, L., Sun, S., Hu, W., Zhang, F., Li, W., 2018. DeepUNet: A Deep Fully Convolutional Network for Pixel-Level Sea-Land Segmentation. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 11 (11), 3954–3962. https://doi.org/10.1109/JSTARS.2018.2833382.

Liu, W., Ma, L., Chen, H.e., Han, Z., Soomro, N.Q., 2017. Sea - Land Segmentation for Panchromatic Remote Sensing Imagery via Integrating Improved MNcut and Chan - Vese Model. IEEE Geosci. Remote Sens. Lett. 14 (12), 2443–2447. https://doi.org/10.1109/LGRS.2017.2768300.

Luo, P., Wang, X., Shao, W., Peng, Z., 2019. Towards understanding regularization in batch normalization. 7th Int Conf. Learn. Represent. ICLR 2019, 1–23.

Ma, L., Liu, Y., Zhang, X., Ye, Y., Yin, G., Johnson, B.A., 2019. Deep learning in remote sensing applications: A meta-analysis and review. ISPRS J. Photogramm. Remote Sens. 152, 166–177. https://doi.org/10.1016/j.isprsjprs.2019.04.015.

Qiao, G., Mi, H., Wang, W., Tong, X., Li, Z., Li, T., Liu, S., Hong, Y., 2018. 55-year (1960–2015) spatiotemporal shoreline change analysis using historical DISP and Landsat time series data in Shanghai. Int. J. Appl. Earth Obs. Geoinf. 68, 238–251. https://doi.org/10.1016/j.jag.2018.02.009.

Ranjbar, S., Zarei, A., Hasanlou, M., Akhoondzadeh, M., Amini, J., Amani, M., 2021. Machine learning inversion approach for soil parameters estimation over vegetated agricultural areas using a combination of water cloud model and calibrated integral equation model. https://doi.org/10.1117/1.JRS.15.018503 15, 018503. https://doi.org/10.1117/1.JRS.15.018503.

Rogers, M.S.J., Bithell, M., Brooks, S.M., Spencer, T., 2021. VEdge_Detector: automated coastal vegetation edge detection using a convolutional neural network. Int. J. Remote Sens. 42 (13), 4805–4835. https://doi.org/10.1080/01431161.2021.1897185.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation. Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics) 9351, 234–241.

Rostami, A., Shah-Hosseini, R., Asgari, S., Zarei, A., Aghdami-Nia, M., Homayouni, S., 2022. Active Fire Detection from Landsat-8 Imagery Using Deep Multiple Kernel Learning. Remote Sens. 14, 992. https://doi.org/10.3390/rs14040992.

Russell, B.C., Torralba, A., Murphy, K.P., Freeman, W.T., 2008. LabelMe: A database and web-based tool for image annotation. Int. J. Comput. Vis. 77 (1-3), 157–173. https://doi.org/10.1007/s11263-007-0090-8.

San, B.T., Ulusar, U.D., 2018. An approach for prediction of shoreline with spatial uncertainty mapping (SLiP-SUM). Int. J. Appl. Earth Obs. Geoinf. 73, 546–554. https://doi.org/10.1016/j.jag.2018.08.005.

Shamsolmoali, P., Zareapoor, M., Wang, R., Zhou, H., Yang, J., 2019. A Novel Deep Structure U-Net for Sea-Land Segmentation in Remote Sensing Images. IEEE J Sel. Top. Appl. Earth Obs. Remote Sens. 12 (9), 3219–3232. https://doi.org/10.1109/JSTARS.2019.2925841.

Shorten, C., Khoshgoftaar, T.M., 2019. A survey on Image Data Augmentation for Deep Learning. J. Big Data 6 (1). https://doi.org/10.1186/s40537-019-0197-0.

van der Werff, H.M.A., 2019. Mapping shoreline indicators on a sandy beach with supervised edge detection of soil moisture differences. Int. J. Appl. Earth Obs. Geoinf. 74, 231–238. https://doi.org/10.1016/j.jag.2018.09.007.

Vos, K., Harley, M.D., Splinter, K.D., Simmons, J.A., Turner, I.L., 2019. Sub-annual to multi-decadal shoreline variability from publicly available satellite imagery. Coast. Eng. 150, 160–174. https://doi.org/10.1016/j.coastaleng.2019.04.004.

Wu, G., de Leeuw, J., Skidmore, A.K., Liu, Y., Prins, H.H.T., 2009. Performance of Landsat TM in ship detection in turbid waters. Int. J. Appl. Earth Obs. Geoinf. 11 (1), 54–61. https://doi.org/10.1016/j.jag.2008.07.001.

Yang, C., Rottensteiner, F., Heipke, C., 2020a. Investigations on skip-connections with an additional cosine similarity loss for land cover classification. ISPRS Ann

Photogramm. Remote Sens. Spat. Inf. Sci. 5, 339–346. https://doi.org/10.5194/isprs-Annals-V-3-2020-339-2020.

Yang, T., Jiang, S., Hong, Z., Zhang, Y., Han, Y., Zhou, R., Wang, J., Yang, S., Tong, X., Kuc, T. yong, 2020. Sea-Land Segmentation Using Deep Learning Techniques for Landsat-8 OLI Imagery. Mar. Geod. 43, 105–133. https://doi.org/10.1080/01490419.2020.1713266.

Yu, X., Wu, X., Luo, C., Ren, P., 2017. Deep learning in remote sensing scene classification: a data augmentation enhanced convolutional neural network framework. GIScience Remote Sens. 54 (5), 741–758. https://doi.org/10.1080/15481603.2017.1323377.

Zarei, A., Hasanlou, M., Mahdianpari, M., 2021. A comparison of machine learning models for soil salinity estimation using multi-spectral earth observation data. ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci. 5, 257–263. https://doi.org/10.5194/isprs-annals-V-3-2021-257-2021.

Zhang, R., Li, G., Wunderlich, T., Wang, L.i., 2021. A survey on deep learning-based precise boundary recovery of semantic segmentation for images and point clouds. Int. J. Appl. Earth Obs. Geoinf. 102, 102411. https://doi.org/10.1016/j.jag.2021.102411.