

Improving the efficiency of using deep learning model to determine shoreline position in high-resolution satellite imagery

Nguyen Thanh Doan^{1*}

¹Moscow State University of Geodesy and Cartography, Moscow, Russia

Abstract. Nowadays, expanding the application of deep learning technology is attracting attention of many researchers in the field of remote sensing. This paper presents methodology of using deep convolutional neural network model to determine the position of shoreline on Sentinel 2 satellite image. The methodology also provides techniques to reduce model retraining while ensuring the accuracy of the results. Methodological evaluation and analysis were conducted in the Mekong Delta region. The results from the study showed that interpolating the input images and calibrating the result thresholds improve accuracy and allow the trained deep learning model to externally test different images. The paper also evaluates the impact of the training dataset on the quality of the results obtained. Suggestions are also given for the number of files in the training dataset, as well as the information used for model training to solve the shoreline detection problem.

Keywords: remote sensing, coastline, shoreline, deep learning, convolutional neural network, deep convolutional neural network, NDVI, NDWI.

1. Introduction

Shoreline and the change in position of this boundary over time are of prime importance to scientists, engineers, and coastal managers. The problem of determining shoreline has long been affecting many spheres of life, because water has always been a valuable resource of mankind. This problem has important implications in areas such as resource management, including water resources, agricultural management, support in the management of emergencies caused by storms, floods, banks erosion of river and sea. Boundary of seas, rivers, and other surface features are important features represented on various types of maps. A shoreline, as defined by the US National Oceanic and Atmospheric Administration, is “The intersection of the land with the water surface. The shoreline represents the line of contact between the land and a selected water elevation. In areas affected by tidal fluctuations, this line of contact is the mean high water line. In confined coastal waters of diminished tidal influence, the mean water level line may be used. The shoreline is defined as MHW” This means that the standard definition of shoreline must be determined using topographic elevation data. This makes it difficult to use remote sensing data to determine the shoreline because individual images do not contain elevation information while stereo satellite images are on-demand and expensive. To overcome this situation, many studies have suggested using indicators on individual images to indicate the shoreline. Indicators differ depending not only on the type of data used, but also on the morphological types of natural coasts in the region of interest. For example, for sandy

* Corresponding author: Nguyen Thanh Doan, email address: ngthanhdogan@outlook.com

beaches and wetlands, the shoreline is vegetation limit; with soft rock cliffs, it is the foot of the hill-slope while for hard rock cliffs, the shoreline is the top of the cliff [1]. However, there is no generally accepted way of defining shoreline indicators. Each researcher can propose to select the most appropriate indicators both for the source data and for the area of their study. [2-5] Because of this inconsistency, it is difficult to use data from different studies and different sources to estimate shoreline changes in a particular area. Lack of synchronization of the input data used leads to the possibility of misjudging shoreline changes, especially in some areas, such as the Mekong Delta in southern Vietnam, where the rate of shoreline change is very high, the accuracy of determining variability becomes even more important. To accurately estimate the change in shoreline, it is necessary to strictly follow the same set of indicators that apply to images.

Currently, research into the application of deep learning technology is a trend that is attracting the attention of many researchers in the field of remote sensing. Deep learning is a set of machine learning algorithms based on artificial neural networks that are used to learn feature at high-level of abstractions using architectures composed of many non-linear transformations. Neural networks have many different architectures to respond to many problems, in which a branch of convolutional neural network architecture was invented with the original task of solving the image classification problem – Convolutional Neural Network (CNN) architecture. Although the theory of convolutional neural network was invented back in the 20th century, this architecture have not received enough attention because they did not initially achieve the accuracy of traditional machine learning methods, due to the lack of large training data and computing power at that time [6]. Only in the last decade, modern convolution neural network models demonstrated outstanding accuracy when performing computer vision tasks. The trend toward reapplication of convolutional networks in scientific research began with the success of the AlexNET model in the ImageNet Large Scale Visual Recognition Challenge on September 30, 2012 (ILSVRC 2012). This network achieved a top-5 error of 15.3%, more than 10.8 percentage points lower than that of the runner up [7]. From the evolution of the architectures, a typical trend is that the networks are getting deeper, e.g., ResNet, which won the champion of ILSVRC 2015, is about 20 times deeper than AlexNet [8]. These aforementioned models were developed to solve the image classification problem, which means that the model will predict which class the inspected image belongs to. The image segmentation problem, also known as the problem of identifying groups of pixels in an image, each characterizing a single semantic object, was solved by deep convolutional network models called U-net by O. Ronneberger and colleagues in 2015 [9]. Many more recent studies have incorporated ResBlock structure from the ResNet model into the U-net architecture to improve image segmentation efficiency, and some studies have used this approach directly to solve the shoreline detection problem, for example, the study of R. Li [10] and Z. Chu [11]. Both above studies used Google Earth image as input data to train the model. Google Earth data has the advantage of ultra-high spatial resolution, but only 3 visible image channels Red, Green and Blue were provided. Remote sensing satellite systems such as Landsat 8, Sentinel 2 often provide data with more image channels, corresponding to more acquired spectral bands, which allows many objects in an image to be identified by the reflection spectrum of that object [12]. Previous studies have not taken advantage of multichannel remote sensing satellite image.

The goal of the study is to apply the deep convolutional network model to determine the shoreline on the Sentinel 2 imagery. In order to achieve this goal, a number of requirements are set:

- Learn to identify the shoreline in satellite image;
- Learn how to create a training dataset from Sentinel 2 imagery;

- Learn how to improve the accuracy of deep convolution model results.

2. Study area

The selected study area is the Mekong Delta located in Vietnam. This is where 19% of the population is concentrated and is the place where 50% of the food for the whole of Vietnam is produced. In some areas of this region, a very fast rate of change of shoreline has been observed, in some places up to more than 20m/year [13,14], therefore, monitoring this change is extremely important to protect natural resources, ensure the safety of residential and agricultural areas. Figure 1 shows an overview map of the Mekong Delta region.

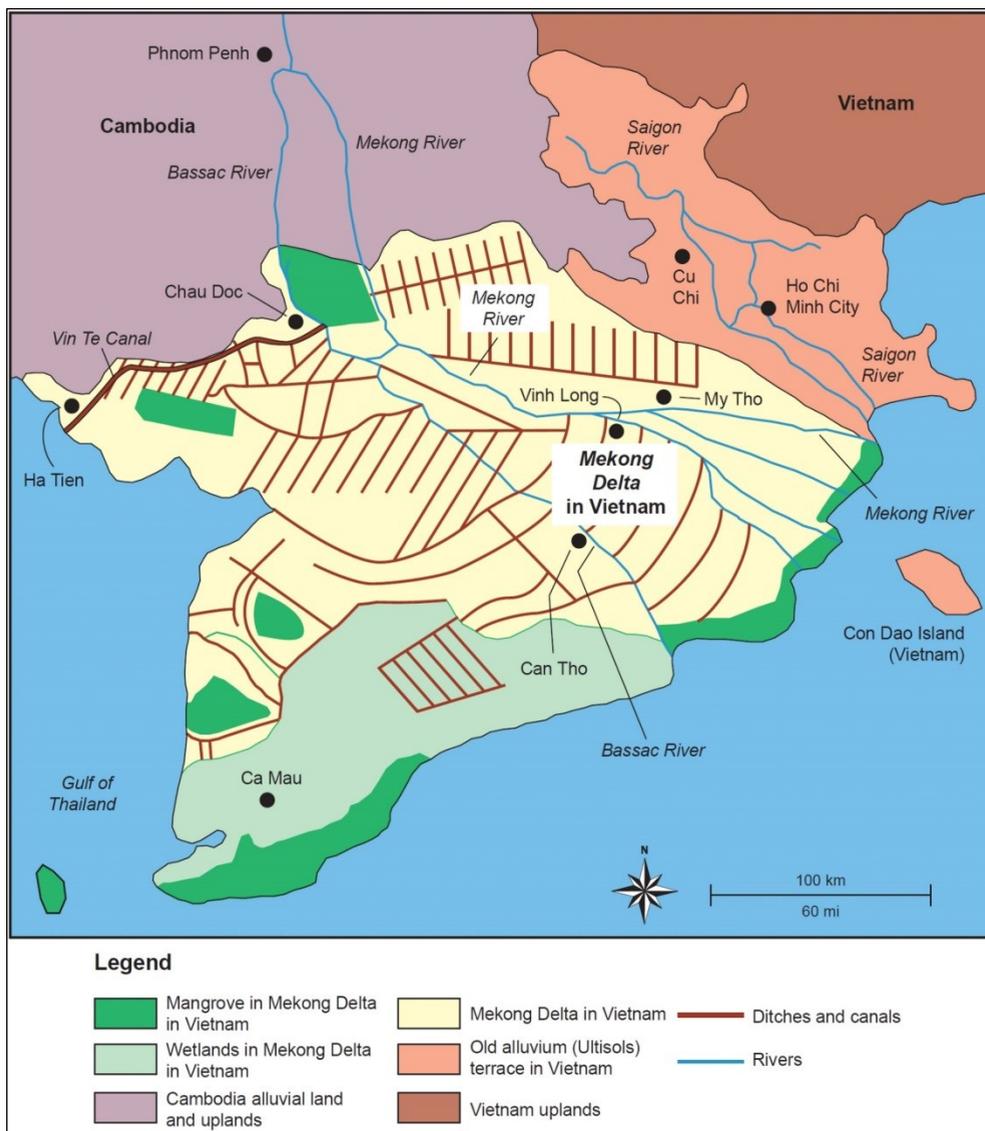


Fig. 1. Map of the Mekong Delta in Vietnam [15]

This area has 4 main morphological types of the coastal zone:

- Coastal mangrove forests - this is a common form of the landscape in Mekong Delta area, concentrated in the south of the delta; [16]
- Sandy coast bounding river mouths - the sandy beach is concentrated at the mouth of the river where the Mekong River flows into the sea;
- Coastal construction - in addition to the two natural coastal forms, people also build coastal constructions for agriculture and human livelihoods, such as sea dams and coastal cities;
- In addition, there are gentle hills and mountains bordering the sea, but this form occupies a small area.

Because coastal areas have many morphological types, using only one indicator for all areas will cause errors when determining shorelines. In this study, author proposes following method for determining the shoreline:

- Using vegetation line as indicator of shoreline when the area bordering the sea is mangrove forest;
- Using wet/dry line as indicator of shoreline when the area bordering the sea is sandy beach;
- If the area bordering the sea coastal construction – using the boundary between the building and the sea as indicator of shoreline.

Figure 2 below is an example of how the shoreline is determined according to the instructions above.



Fig. 2. Examples of indicator selection

3. Methodology

In general, the technique of using a convolutional neural network model to be used in any field of science has to go through three steps.

1. The first step is to create training dataset.
2. The second step is to train the model from the prepared dataset.
3. The third step is to use the trained model to external test.

Since the structure of the neural network is constant during and after training, increasing the efficiency of using the deep learning model is the optimization of the first and third step.

3.1 Dataset creation and training model

Unlike traditional machine learning methods, the convolutional neural network model uses a training dataset consisting of input images and masks corresponding to these images. In the process of training the model, training dataset will be used to optimize model's parameters so that the model prediction results are closest to the given mask. A common deep convolutional neural network model for solving image segmentation problems often has ten million to hundreds of millions of such parameters. Therefore, to train a convolutional neural network model, it is necessary to use a large amount of input data. On the other hand, the whole process of training a deep learning model is automatic, the parameters of the model are independent of the person, so the quality of the input data directly affects the quality of the model. In other words, to get a good model, the input must also be of good quality.

The process of creating a training dataset consists of two steps: selecting the images to be used for training and creating appropriate masks. The entire process of creating data and training the neural network model was described by flowchart in Figure 3 below.

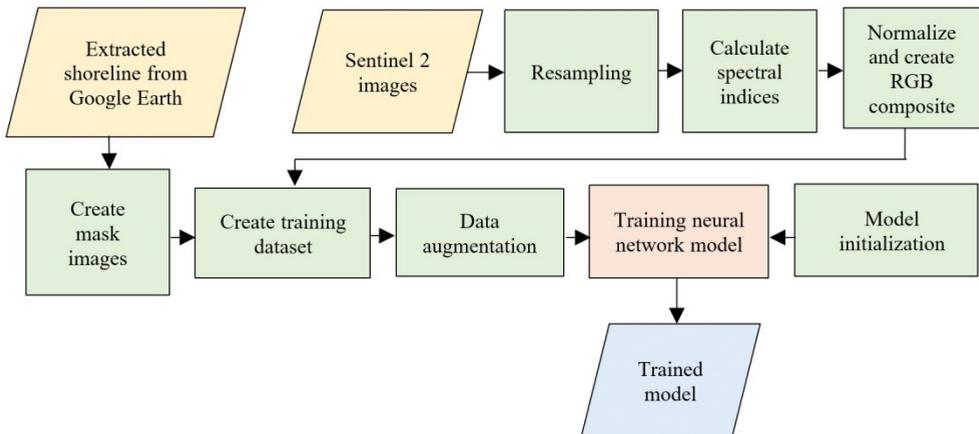


Fig. 3. The process of creating data and training convolutional neural network model for shoreline determination in high-resolution satellite imagery.

Sentinel-2 is an Earth observation mission from the Copernicus Programme that systematically acquires optical imagery at high spatial resolution (10 m to 60 m) over land and coastal waters. 4 of Sentinel 2 bands have 10m spatial resolution (Blue band, Green band, Red band and Near Infrared band) [17]. Sentinel 2 data has been available since 2015, so it can be used to estimate shoreline changes in the study area from 2015 to the present.

Shoreline is a dynamic feature, so it is extremely difficult to find sample data to create a large number of masks. Shoreline data from field measurements using GPS receivers are not available, so in this study author suggests using a shoreline extracted from Google Earth's ultra-high resolution satellite imagery. Google's partners such as Maxar and CNES/Airbus provide these ultra-high resolution satellite images for Google Earth across most countries. The advantage of this image source in addition to the high resolution is the information about the image's sensing date. This is a great advantage, especially important for the task of defining a shoreline that is constantly changing. Knowing the date of Google Earth data collection, you can select Sentinel 2 images taken in the corresponding time period to create a training dataset. The closer the Sentinel 2 image's sensing date is to the Google Earth image's sensing date, the more reliable the training data will be. Therefore,

author proposes to use the Sentinel data 2 and the data from Google Earth with difference in the sensing date of is not more than 1 month. Shoreline will be determined manually on Google Earth software, then exported as vector feature. This vector data will then be converted into a mask image to create a pair with the data from the Sentinel 2 image. Mask image and Sentinel 2 image used for training must have the same resolution. Since the spatial resolution of Sentinel 2 images is only 10m maximum, converting vector data to mask image with equivalent resolution will lose a lot of original information. In order to make better use of the sample data, the author recommends upsampling Sentinel 2 images to a resolution of 2.5m and then conducting training. The upscaling method chosen is bilinear interpolation.

Spectral characteristics must be extracted from the Sentinel 2 image channels in order to isolate the three indicators suggested above before using it as a dataset. There are two spectral indices that are widely used - Normalized Difference Vegetation Index (NDVI) [18] and Normalized Difference Water Index (NDWI) [19].

Normalized Difference Vegetation Index (NDVI) is an index used to highlight vegetation objects on remote sensing images. This index is determined by the red channel (R) and the near infrared channel (NIR) according to the following formula:

$$NDVI=(NIR-R)/(NIR+R) \quad (1)$$

Normalized Difference Water Index (NDWI) is an index used to highlight surface water bodies in remote sensing images. This index is determined by the green channel (G) and the near infrared (NIR) channel using the following formula:

$$NDWI=(G - NIR)/(G + NIR) \quad (2)$$

In addition to the above two indicators, the author suggests adding a red channel (R) to make wet sand and dry sand more visible in the image.

Another problem that should be noted is that because the input data is taken from many images, from different areas and from different times, therefore, the reflection properties in each image are not the same, which causes difficulties both in training the model and in using it. Therefore, before training the model, it is necessary to normalize the input data. Since two indices NDVI and NDWI are normalized indices, we only need to convert the values of these two indices from [-1,1] to [0,255]. The red channel of the Sentinel 2 image will be normalized differently depending on each frame so that the sandy object is clearly visible and the frame is not underexposed, usually the digital number value range [0,4000] is converted to [0,255]. The three resulting channels will be composited into a false-color RGB composite image. Because the convolutional neural network model requires a large amount of GPU memory when running training, usually a Sentinel 2 image is split into many smaller tiles for processing, rather than processing the entire image at the same time. Figure 4 shows an example of cropped images in the training dataset.

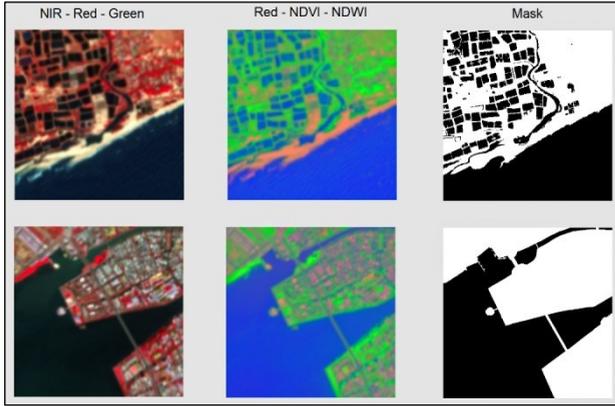


Fig. 4. Example of training data

This study applies an improved neural network model from the U-Net architecture, which was introduced by Olaf Ronneberger and his colleagues in 2015 [9]. The U-Net architecture is based on a fully convolutional network and has been modified to provide better segmentation of medical images. This neural network structure consists of two parts, the encoder and the decoder, which are built symmetrically with each other. The encoding part represents itself as a typical convolutional classification network architecture that is responsible for classifying objects on the image into the required classes. The decoding part of this architecture is also a convolutional network but has a reverse structure of the encoder, which is responsible for determining which class each pixel on the image belongs to. They are built with a mirror-symmetrical structure and skip joins have been added between the symmetric layers using the concatenation operator to ensure that the results obtained with the highest detail [7]. This study uses the ResNet34 convolutional network structure to create encoder and decoder. Adding skip connections using the plus operator right in the layers of the encoder allows the Resnet model structure to reach many layers without negatively affecting the result. The structure of the model used in this study is shown in Figure 5.

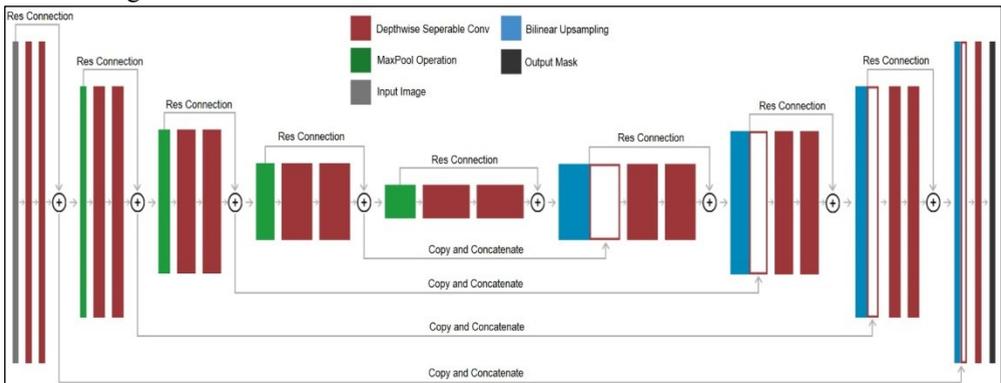


Fig. 5. The structure of the deep convolutional neural model used in this study

Some studies have used a similar model architecture to solve remote sensing problems, for example, in the studies of Ruirui Li [10] and Foivos I.Diakogiannis [20]. The process from creating dataset to training model is carried out automatically in Python environment with the help of FastAI and Solaris libraries. Currently, python is a programming environment used by many researchers to research the application of artificial intelligence in general and deep learning in particular due to the support of strong machine learning and deep learning

libraries. This approach makes the process of using an convolutional neural network as flexible as possible. This allows users to easily create and customize neural network structures to solve a research problem without relying on third-party software.

3.2 Improving the efficiency of using deep learning model to determine the shoreline on the image

Once the trained model is obtained, it will be used to define the shoreline on the satellite image. The first step in this process is similar to the creation of the training dataset, except that there is no mask creation step. The model relies on inputs to generate predictions in the form of an image where the digital number (DN) fluctuating in the range [0,1] of each pixel represents the probability that the pixel is a water surface or land surface. It can be said that using a convolutional neural network model is a fuzzy classification method to identify objects on remote sensing satellite images. Processing is performed according to the flowchart in Figure 6 below.

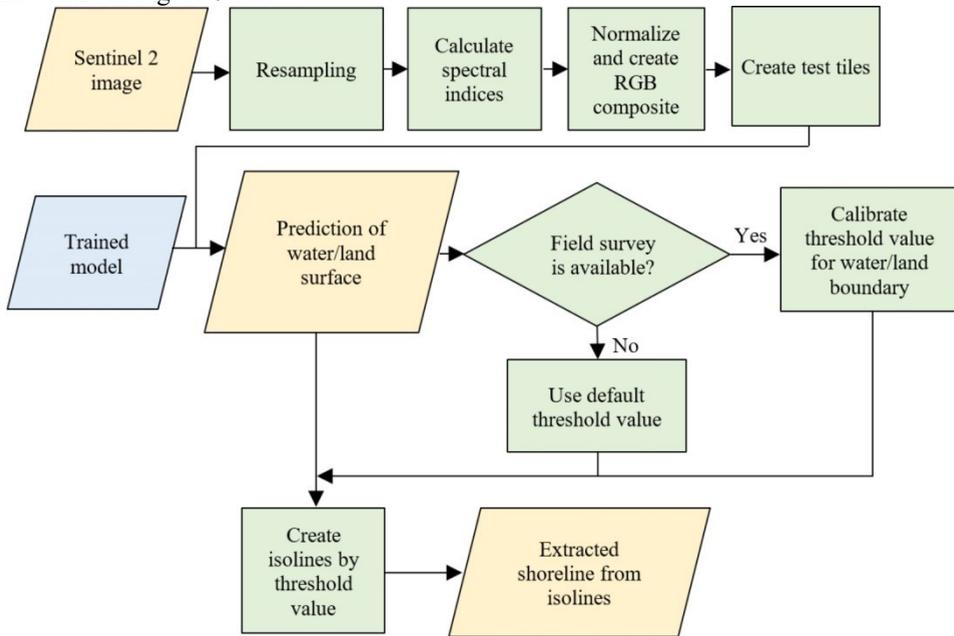


Fig. 6. The process of identifying shorelines on Sentinel 2 image using convolutional neural network model

As shown in Figure 6, the shoreline is not defined directly by the boundary between pixels, but by the subpixel isoline passing through the threshold value. This technique is proposed for two purposes. The first purpose is to increase the accuracy of the received shoreline by threshold calibration. The second purpose is to help the resulting shoreline have a curve shape similar to a real shoreline, rather than a jagged boundary between pixels.

By default, 0.5 can be set as the threshold value for determining the water/land boundary. However, since the reflectivity of the objects in each image is different, the threshold must be recalibrated to ensure the most accurate results obtained. In this process, a given area in the image is used as a sampling area, shoreline position of which is known from the results of field measurements. The threshold value is determined by the average DN value of the prediction results by the convolutional neural network model on the known shoreline. The threshold calibration process allows a trained model to be used with multiple images captured at different times and areas, even from different image sources, without the need for retraining. In fact, training an convolutional neural network model takes much longer

than the execution time to obtain water/land surface classification results. Therefore, limiting the need to retrain the neural network model will save a lot of time and effort and still give acceptable results. Determining the shoreline from the contour line will help the result not depend on the boundary of each pixel on the image, in other words, this is a method that can obtain the shoreline with subpixel accuracy.

3.3 Criteria for evaluating the error of results

The process of evaluating the accuracy of the shoreline causes a problem requiring revision evaluation method. Research by Amy H. Pickens shows that only the area that near the land/water boundary causes classification error, the area far from this boundary has a classification error close to 0 [21]. This leads to the fact that when studying a specific coastal area, simply changing the area of interest towards increasing/decreasing sea or land area will lead to an error in accuracy if the accuracy is determined by the traditional coefficients, such as the Intersection over Union (IoU) metric [22].

Shoreline on the other hand is a line object on the map. According to Geospatial Positioning Accuracy Standards [23], horizontal error of a point i is defines as is defined as $\sqrt{(x_{data, i} - x_{check, i})^2 + (y_{data, i} - y_{check, i})^2}$. Therefore, the resulting shoreline by classification process must also be determined by statistical evaluation of the distance between it and the shoreline measured in the field. This approach allows to get the accuracy not only not depending on the region of interest, but also to determine the accuracy in accordance with the requirements when creating a map. According to United States National Map Accuracy Standards, For maps on publication scales larger than 1:20,000, not more than 10 percent of the points tested shall be in error by more than 1/30 inch, measured on the publication scale [24]. This shows that if the obtained shoreline has a 90% of points have an error of less than 8.467 m, the results can be used to create a 1:10000 scale map.

4. Experiments and analysis

The experiment was carried out on a computer with the configuration shown in Table 1. The operating system installed on this computer is Ubuntu 20.04. Satellite images are processed using QGIS 3.16. The process of initializing and operating the convolutional neural network is done in the python 3 programming environment with the necessary support libraries including FastAI, Pytorch, Solaris, etc.

Table 1. Experimental enviroments.

CPU	Intel(R) Core(TM) i7-7700HQ 2.80GHz
GPU	GeForce GTX1060 6GB
RAM	16GB
Hard disk	Samsung EVO 970 1TB
Operating system	Ubuntu 20.04

The process of creating sample data uses 6 different Sentinel 2 images, covering the central and southern regions of Vietnam, respectively, 6 shoreline samples were defined as polygons on Google Earth. The result is a dataset of 600 image-mask pairs used to train the model. Specific information is presented in Table 2 below.

Table 2. Input data

Name of Sentinel 2 image	Sensing date of Sentinel 2 image	Sensing date of Google Earth data	Morphology type of coast zone
S2A_MSIL1C_20190217T030801_N0207_R075_T48PYS_20190217T074141	Feb 17th, 2019	Feb 11th, 2019	Sandy beach
S2A_MSIL1C_20190312T031541_N0207_R118_T48PVS_20190312T074859	Mar 12th, 2019	Mar 04th, 2019	Narrow strip of mangrove forest interspersed with residential areas
S2A_MSIL2A_20190627T030551_N0212_R075_T48PZC_20190627T071355	Jun 27th, 2019	Jun 26th, 2019	City with high construction density, steep hills and sandy beaches
S2B_MSIL1C_20191202T032109_N0208_R118_T48PVS_20191202T065246	Dec 02nd, 2019	Dec 30nd, 2019	Narrow strip of mangroves interspersed with residential areas
S2A_MSIL1C_20200113T031101_N0208_R075_T48PYS_20200113T061224	Jan 13th, 2020	Jan 14th, 2020	Mangrove forest
S2A_MSIL1C_20200205T031921_N0209_R118_T48PVR_20200205T061921	Feb 05th, 2020	Jan 23rd, 2020	Narrow strip of mangrove forest

To study influence of quantity and quality of input data on the quality of the model, we gradually subdivided the training dataset and trained the model corresponding to that amount of data. In parallel, creating datasets with test image data is false color composite NIR-Red-Green to study the influence of data types used to train the model. Table 3 shows how to prepare dataset in this study.

Table 3. Variations of dataset

	False color composite	
	Red-NDVI-NDWI	NIR-Red-Green
Amount of data	600 image-mask pairs	600 image-mask pairs
	300 image-mask pairs	300 image-mask pairs
	150 image-mask pairs	150 image-mask pairs
	50 image-mask pairs	50 image-mask pairs
	20 image-mask pairs	20 image-mask pairs
	10 image-mask pairs	10 image-mask pairs

The above process is built in the form of an inverted pyramid. The largest dataset containing information both in the land area, at the water-land border, as well as in the sea area. Smaller datasets lack images that related only to the sea or land. In other words, smaller datasets will contain more useful information than larger datasets. However, a larger dataset has the advantage that it contains a wide variety of information and is expected to give better results because the model uses more data for training.

Model training is done in python programming environment. The model is evaluated for accuracy automatically by F1-score. However, this evaluation process is only performed on the imported dataset, so it is not possible to evaluate the F1-score of one dataset with the F1-score of another dataset. The F1-score value of the datasets in this study ranges from 0.95 to 0.97. The model is trained many epochs until the F1-score fluctuates not significantly. Figure 7 shows how the F1-score value changes after epochs.

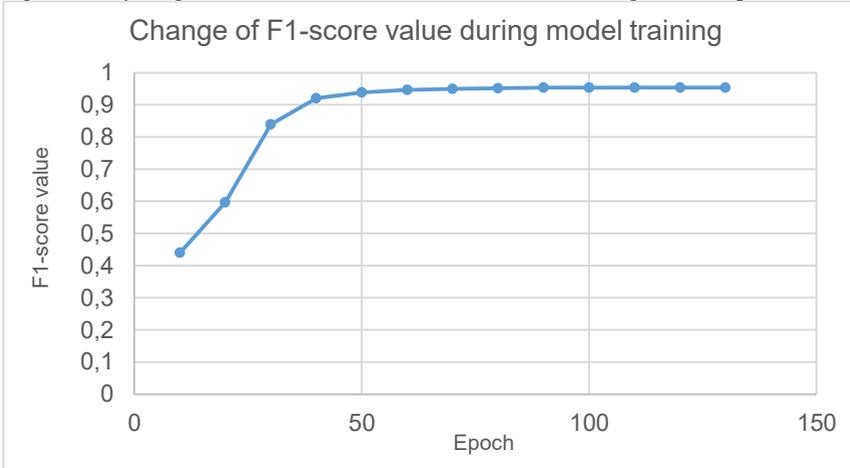


Fig. 7. Change of F1-score value during model training (dataset of 10 image-mask pairs NIR-Red-Green combination)

In addition, the shoreline is extracted by isolines, so it is impractical to estimate accuracy by pixels. The accuracy of the results was determined by horizontal accuracy as mentioned in section 3.3. To evaluate the quality of the models after training, these resulting models will be used to define the shoreline in an area with a known position of shoreline in the west area of Mekong delta. The image used is a Sentinel 2 image captured on February 20th, 2020. The area used to determine the threshold value is an area to the north of the AOI, which does not overlap with the area used to evaluate the results, indicated by the dashed black line (see Fig. 7).

For comparison, this study also used Support Vector Machine (SVM) classifier in ENVI 5 to evaluate the results when using NDVI-NDWI-Red composite images without additional other enhancements. The results are shown in Figure 8 below.

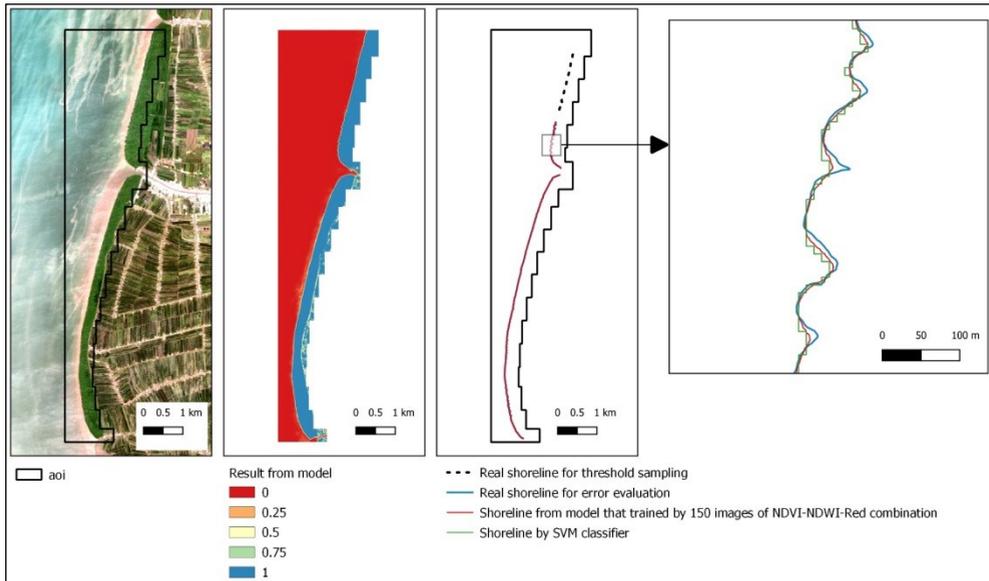


Fig. 8. Result obtained from the model that trained by 150 images of NDVI-NDWI-Red combination

Table 4 below shows the result of the error estimation process for neural network models at test area:

Table 4. Error estimation

Dataset		Mean of distance (m)	Standard deviation of distance (m)	Max distance of 90% points (m)
NDVI-NDWI-Red combination	10 pairs	3.233	2.674	7.104
	20 pairs	3.144	2.447	6.565
	50 pairs	2.756	1.853	5.245
	150 pairs	2.265	1.654	4.413
	300 pairs	3.523	2.111	6.374
	600 pairs	3.281	2.256	6.407
NIR-Red-Green combination	10 pairs	5.286	3.780	10.600
	20 pairs	4.379	3.230	8.957
	50 pairs	4.463	4.155	9.020
	150 pairs	4.218	5.051	8.653
	300 pairs	3.003	2.044	5.840

	600 pairs	3.463	2.397	6.712
Method SVM		3.415	2.985	6.944

The table shows that the NDVI-NDWI-Red combination gives horizontal accuracy from 4.413m to 7.104m, while the NIR-Red-Green combination gives horizontal accuracy from 5.840m to 10.600m. This shows that using a combination of optical indices improves the quality of results when performing external tests on new images.

When comparing results between models using different amounts of training data, we can see that the trend of both combinations is that model accuracy is lowest when using 10 image-mask pairs, reaching 7.104m with the NDVI-NDWI-Red combination and 10.600m with the NIR-Red-Green combination. Horizontal accuracy gradually increases with increasing input data, reaching 4.413m with 150 pairs of NDVI-NDWI-Red combination and 5.840m with 300 pairs of NIR-Red-Green combination. Then, increasing the amount of training data no longer increases the accuracy of the model.

When comparing the results of the deep convolutional network method with the SVM method, the deep convolutional network method still gives a higher horizontal accuracy, reaching 4.413m compared to the 6.944m accuracy of SVM method.

5. Discussion and conclusions

This article presents techniques for using a deep learning model to determine the position of a shoreline in high-resolution satellite imagery. To achieve this goal, the following tasks were solved:

- Identify the study area;
- Agree on how to use indicators to determine the shoreline on remote sensing images corresponding to the study area;
- Propose a method to increase the efficiency of creating training dataset for that indicator;
- Propose a shoreline determination method based on a modern deep convolutional neural network model and ensure that the model works with different images, in different areas.
- Recommend an appropriate way to evaluate the accuracy of shoreline;
- Testing and evaluating the results.

Based on the results, the amount of initial data used for training convolutional neural network model does not need to be too large, it is important that the data used for training reflects characteristics of the object that needs to be classified in the image.

By interpolating the input image and calibrating the threshold for shoreline contouring, the convolutional neural network model can use the Sentinel 2 image at 10m resolution to extract the shoreline with horizontal accuracy of 4.413m. The results obtained with a deep learning model trained on pairs of NDVI-NDWI-Red images with the above error can be used to create a map at a scale of 1:10000.

In conclusion, NDVI-NDWI-Red combination is at least 24% efficient compared to using only conventional NIR-Red-Green channels of Sentinel 2 image. Comparison with the support vector machine also shows that the deep learning model gives result's accuracy

36% higher in the study area. The method presented in this paper can be studied to apply to many other objects besides shoreline determining.

This study outputs the shoreline as a subpixel line, so the process of estimating accuracy is problematic. However, in the author's opinion, the use of horizontal accuracy is a reasonable criterion for evaluating the results obtained. It meets national map accuracy standards, so for any project that uses this metric, it is also convenient to calculate which inputs meet project requirements.

However, there are still some aspects that need to be studied more concretely, such as how each shoreline-facing object type affects the threshold value, quantifying the effect of AOI selection on threshold variation. Variation in traditional indicators such as F1-estimation, IoU, and kappa coefficient.

References

1. A. Puissant, S. Lefèvre, J. Weber, *Coastline extraction in VHR imagery using mathematical morphology with spatial and spectral knowledge*, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. **XXXVII**, Part B8, pp. 1305-1310, Beijing, China (2008)
2. E. Boak, I. Turner, *Shoreline Definition and Detection: A Review*, J. of Coastal Research, **21**, pp. 688-703, (2005)
3. V. Baiocchi, R. Brigante, D. Dominici, F. Radicioni, *Coastline Detection Using High Resolution Multispectral Satellite Images*, FIG Working Week 2012, 6-10 May 2012, Rome, Italy, (2012)
4. P. Maglione, C. Parente, A. Vallario, *Coastline extraction using high resolution WorldView-2 satellite imagery*, European Journal of Remote Sensing, **47:1**, pp. 685-699, (2014)
5. K. Smith, J. Terrano, J. Pitchford, M. Archer, *Coastal Wetland Shoreline Change Monitoring: A Comparison of Shorelines from High-Resolution WorldView Satellite Imagery, Aerial Imagery, and Field Surveys*, Remote Sensing, **13**, 3030, (2021)
6. J. Gu, Zh. Wang, J. Kuen, L. Ma, A. Shahroudy, B. Shuai, T. Liu, X. Wang, G. Wang, J. Cai, T. Chen, *Recent advances in convolutional neural networks*, Pattern Recognition, **77**, pp. 354-377, (2018)
7. A. Krizhevsky, I. Sutskever, G. Hinton, *ImageNet Classification with Deep Convolutional Neural Networks*, Communications of the ACM, **60**, 6, pp. 84-90, (2017)
8. K. He, X. Zhang, S. Ren, J. Sun, *Deep Residual Learning for Image Recognition*, 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770-778, (2016)
9. O. Ronneberger, P. Fischer, T. Brox, Thomas, *U-Net: Convolutional Networks for Biomedical Image Segmentation*, Lecture Notes in Computer Science, **9351**, pp. 234-241, (2015)
10. R. Li, W. Liu, L. Yang, S. Sun, W. Hu, F. Zhang, W. Li, *DeepUNet: A Deep Fully Convolutional Network for Pixel-Level Sea-Land Segmentation*, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, **11**, 11, pp. 3954-3962, (2018)

11. Z. Chu, T. Tian, R. Feng, L. Wang, *Sea-Land Segmentation With Res-UNet And Fully Connected CRF*, 2019 IEEE International Geoscience and Remote Sensing Symposium, pp. 3840-3843, (2019)
12. M. A. Ridwan, N.A.M. Radzi, W.S.H.M.W. Ahmad, I.S. Mustafa, N.M. Din, Y.E. Jalil, A.M. Isa, N.S. Othman, W.M.D.W. Zakiet, *Applications of Landsat-8 Data: a Survey*, International Journal of Engineering & Technology, **7**, 4.35, pp. 436-441, (2018)
13. M. Besset, N. Gratiot, E. J. Anthony, F. Bouchette, M. Goichot, P. Marchesiello, *Mangroves and shoreline erosion in the Mekong River delta, Viet Nam*, Estuarine, Coastal and Shelf Science, **226**, 106263, (2019)
14. B. K. Veettil, X. Q. Ngo, T. T. T. Ngo, *Changes in mangrove vegetation, aquaculture and paddy cultivation in the Mekong Delta: A study from Ben Tre Province, southern Vietnam*, Estuarine, Coastal and Shelf Science, **226**, 106273, (2019)
15. K. R. Olson, L. W. Morton, *Polders, dikes, canals, rice, and aquaculture in the Mekong Delta*, Journal of Soil and Water Conservation, **73**, 4, (2018)
16. T. D. Truong, L. H. Do, *Mangrove forests and aquaculture in the Mekong river delta*, Land Use Policy, **73**, pp. 20-28, (2018)
17. Sentinel-2 Products Specification Document, pp. 35-36, (2021)
18. Zh. Jiang, A.R. Huete, J. Chen, Y. Chen, J. Li, G.n Yan, X. Zhang, *Analysis of NDVI and scaled difference vegetation index retrievals of vegetation fraction*, Remote Sensing of Environment, **101**, 3, pp. 366-378, (2006)
19. S. K. McFeeters, *The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features*, International Journal of Remote Sensing, **17**:7, pp. 1425-1432, (1996)
20. F. I. Diakogiannis, F. Waldner, P. Caccetta, C. Wu, *ResUNet-a: A deep learning framework for semantic segmentation of remotely sensed data*, ISPRS Journal of Photogrammetry and Remote Sensing, **162**, pp. 94-114, (2020)
21. A. H. Pickens, M. C. Hansen, M. Hancher, S. V. Stehman, A. Tyukavina, P. Potapov, B. Marroquin, Z. Sherani, *Mapping and sampling to characterize global inland water dynamics from 1999 to 2018 with full Landsat time-series*, Remote Sensing of Environment, **243**, 117792, (2020)
22. H. Rezatofighi, N. Tsoi, J. Gwak, A. Sadeghian, I. Reid, Ian, S. Savarese, *Generalized Intersection over Union: A Metric and A Loss for Bounding Box Regression*, 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 658-666. (2019)
23. Geospatial Positioning Accuracy Standards Part 3: National Standard for Spatial Data Accuracy, pp. 3-10. (1998)
24. Map Accuracy Standards, USGS Fact Sheet 171-99, (1999)