





Automatic Detection of Waterbodies from Satellite Images Using DeepLabV3+

Seifedine Kadry¹ , Mohammed Azmi Al-Betar², Sahar Yassine¹,
Ramya Mohan³, Rama Arunmozhi³, and Venkatesan Rajinikanth³ 

¹ Department of Applied Data Science, Noroff University College, 4612 Kristiansand, Norway
skadry@gmail.com, sahar.yassine@noroff.no

² Artificial Intelligence Research Center (AIRC), College of Engineering and Information
Technology, Ajman University, Ajman, United Arab Emirates

³ Department of Computer Science and Engineering, Division of Research and Innovation,
Saveetha School of Engineering, SIMATS, Chennai 602105, India

Abstract. In recent times, it has become increasingly popular to examine a wide range of environmental and earth data using remote sensing schemes supported by satellite images (SI). Due to the complex nature of spatial, spectral, and temporal characteristics of SI, it is quite difficult for automatic analysis to be performed as it requires specially designed algorithms. As part of the research proposal, an enhanced deep-learning scheme will be implemented in order to extract the waterbodies from the chosen SIs. The phases involved in this scheme includes; (i) the collection and resizing of images, (ii) Shannon's Entropy preprocessing, (iii) DeepLabV3+ extraction of waterbodies, (iv) the comparison of extracted sections with ground truth (GT) and the calculation of performance metrics, and (v) validation of the effectiveness of the implemented scheme. In the proposed work, the proposed multi-thresholding approach is combined with DeepLabV3+ in order to get the waterbodies to be extracted from the chosen test images. DeepLabV3 is demonstrated to have excellent segmentation performance when it is compared to UNet and SegNet. The experimental results of this scheme indicate that DeepLabV3+ results in higher Jaccard value (>89%), Dice value (>93%), and segmentation accuracy value (>97%) when compared to other pretrained segmentation schemes.

Keywords: Earth data · Satellite image · Waterbodies · DeepLabV3+ · Segmentation

1 Introduction

The rapid improvement in technology and the availability of various computing facilities helped to implement the Artificial Intelligence (AI) schemes to analyse a variety of the database. Examination of the satellite data is one of the common procedure in geoscience studies and the outcome of this assessment helps to provide a variety of vital information regarding various key geoscience information for necessary monitoring and decision making tasks [1, 2].

Detection of waterbodies from the Satellite Imagery (SI) is one of the key research area and this helps to achieve the following;

- Water resource management: Accurate mapping of waterbodies is important for the management of water resources such as rivers, lakes, and reservoirs. It helps in monitoring changes in water levels, identifying areas of water scarcity or abundance, and planning for the allocation of water resources.
- Flood monitoring: Waterbody detection is crucial in monitoring floods, which can cause significant damage to property and human lives.
- Environmental studies: It can help in the identification of wetlands, which are critical habitats for wildlife and provide several ecosystem services such as water purification and carbon sequestration.
- Climate change studies: Changes in waterbodies, such as shrinking of glaciers, melting of sea ice, and changes in river flow patterns, can provide insights into the impacts of climate change. Detection of these changes can help in developing mitigation and adaptation strategies.
- Navigation and transportation: It helps in the planning of shipping routes, construction of ports, and monitoring of maritime activities.

In the literature, there exists a variety of procedures to detect the waterbodies from SIs, such as image segmentation, thresholding, and machine learning algorithms. Compared to the traditional image processing schemes, the AI supported methods are widely adopted in the literature for automatic examination and assesment of the SI.

Convolutional neural networks (CNNs) are a type of deep learning algorithm that has shown great promise in waterbodies analysis from satellite imagery. CNNs are well-suited for image analysis tasks as they can automatically learn and extract relevant features from images without the need for manual feature engineering. CNNs can be trained using labeled satellite images to detect and classify waterbodies at different scales. The labeled images are typically manually annotated by experts to indicate the location of waterbodies. The CNN then learns to recognize patterns in the labeled images and applies them to new images to detect waterbodies [3, 4].

The CNNs are a powerful tool for waterbodies analysis from satellite imagery. They can be used for detection and classification of coastal waterbodies, mapping of inland waterbodies, flood monitoring, and detection of water quality parameters. The information obtained from CNNs can be used for various applications such as water resource management, environmental studies, and disaster response planning. In the proposed research, DeepLabV3+ is considered to extract and evaluate the waterbodies from the choen SI. The various stages of this scheme involves in; image collection and resizing, preprocessing with Shannon's Entropy (SE), DeepLabV3+ supported segmentation and computing the necessary performance measures by comparing the mined region with ground truth (GT). The experimental outome of this study confirms that the proposed scheme is efficient in achiving a better segmentation outcome on the chosen image database.

This work includes of five sections; Sects. 2 and 3 demonstrates the literature review and methodology, Sects. 4 and 5 present experimental outcomeand conclusion of this study.

2 Literature Review

Recently, deep-learning based examination of SIs are widely discussed by the researchers. The work of Zhang et al. (2021) proposes a method for waterbodies segmentation using deep convolutional neural networks (DCNNs). The proposed method achieved high accuracy in waterbodies segmentation in high-resolution satellite images [5]. The study of Gao et al. (2022) proposes an unsupervised method for waterbodies segmentation using generative adversarial networks (GANs) and region merging. The proposed method achieved high accuracy in waterbodies segmentation in Landsat 8 satellite images [6]. The research by Yuan et al. (2020) proposes a deep learning-based method for waterbodies segmentation using data augmentation techniques such as rotation, flipping, and scaling. The proposed method achieved high accuracy in waterbodies segmentation in multispectral satellite images [7].

The work of Rana and Babu (2022) compares the performance of different waterbodies segmentation techniques, including thresholding, segmentation using morphological operations, and deep learning-based methods [8]. The review work by Babu and Rajam (2020) provides an overview of different machine learning techniques for waterbodies segmentation from satellite images, including thresholding, clustering, and deep learning-based methods [9]. The paper also discusses the challenges and future directions in waterbodies segmentation. Based on this motivation, proposed research also implements automatic evaluation of waterbodies from the chosen satellite images.

To achieve a better segmentation accuracy, this research considers the preprocessed SIs and implements the DeepLabV3+. The experimental results of this work confirms that the proposed scheme works well on the chosen images.

3 Methodology

This section of the research presents the methodology proposed in this research to examine the SIs. Figure 1 depicts the various stages existing in the waterbodies detection scheme with DeepLabV3+.

The necessary SI database is collected from [10] and then every image and the GT are resized to $512 \times 512 \times 3$ pixels. These images are then enhanced using the Shannon's entropy supported tri-level thresholding along with the Mayfly-Algorithm (SE + MA) [10]. The waterbodies region from the enhanced image is then mined using the DeepLabV3+ and the extracted section is compared with the Ground-Truth (GT) and the necessary performance metrics are computed. Based on the mean value of these measures, the merit of the proposed technique is verified.

3.1 Database

The necessary test images for this study is obtained from [10]. The considered image and the ground-truth (GT) is resized to 512×512 pixels and in this work a total of 1000 images are considered for the examination. Among the considered images, 80% of the database is chosen for training the pretrained segmentation system and 20% is considered for testing the performance. The sample test image and its related GT is depicted in Fig. 2.

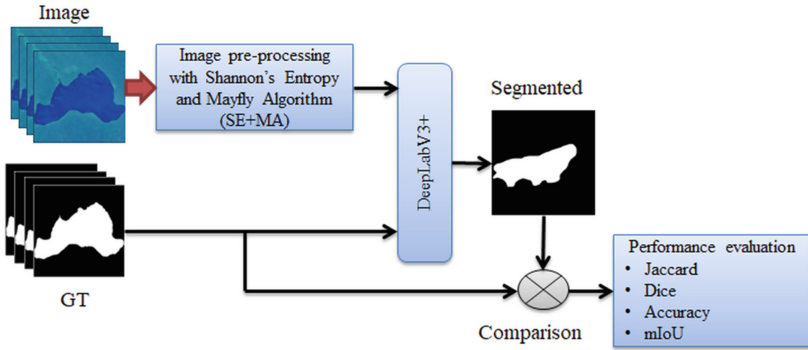


Fig. 1. Proposed satellite image evaluation approach

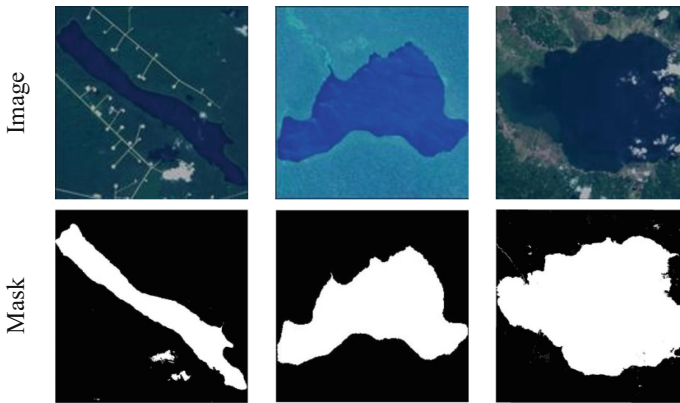


Fig. 2. Benchmark satellite image database

3.2 Thresholding

The chosen RGB scaled test images are preprocessed with a tri-level thresholding to enhance the visibility of the waterbody section for better segmentation. The related thresholding considered in this work can be found in [11–13].

Shannon's Entropy (SE) thresholding is a technique used in image processing to automatically determine a threshold value for image segmentation. This technique uses SE formula to calculate the optimal threshold value for a given image. The basic idea behind SE thresholding is that the threshold value should be chosen in such a way that it maximizes the difference in entropy between the object and background regions of the image. The entropy of an image can be calculated using the formula shown in Eq. (1):

$$H = - \sum (p_i * \log_2(p_i)) \quad (1)$$

where p_i is the probability of the intensity level i occurring in the image. The sum is taken over all possible intensity levels in the image.

To apply Shannon's entropy thresholding, we first calculate the entropy of the entire image. We then consider all possible threshold values between the minimum and maximum intensity values in the image. For each threshold value, we calculate the entropy of the object and background regions of the image separately. The threshold value that maximizes the difference between the entropies of the object and background regions is then selected as the optimal threshold.

Once the optimal threshold value has been determined, the image can be segmented into object and background regions by assigning all pixels with intensity values above the threshold to the object region, and all pixels with intensity values below the threshold to the background region. Shannon's entropy thresholding is a simple and effective technique for image segmentation. However, it is important to note that the technique may not work well for images with non-uniform backgrounds or objects with complex shapes. In such cases, more sophisticated thresholding techniques may be required. The optimal thresholding with SE is obtained using the Mayfly Algorithm (MA) and the complete information regarding this scheme can be found in [14, 15].

3.3 DeepLabV3+

DeepLabV3+ is a state-of-the-art deep neural network architecture for semantic segmentation, which was proposed by Zhang et al. in 2020 [16]. It is an extension of the previous DeepLabV3 architecture, which was designed to address the limitations of earlier semantic segmentation models, such as low spatial resolution, coarse object boundaries, and over-segmentation.

The core idea of DeepLabV3+ is to integrate the features extracted by a deep convolutional neural network (CNN) with a powerful spatial pyramid pooling module and a decoder network that restores the spatial resolution of the output segmentation map. The model is built on top of a pre-trained backbone network, such as ResNet or Xception, which is used to extract high-level features from the input image.

The spatial pyramid pooling module in DeepLabV3+ is designed to capture features at multiple scales and at different levels of granularity. This module takes the output feature maps from the backbone network and applies pooling operations at different scales, with each scale capturing features at a different level of detail. This allows the model to capture both global and local context information and better distinguish between objects that are close together.

The decoder network in DeepLabV3+ is used to restore the spatial resolution of the output segmentation map, which is reduced by the pooling operations in the spatial pyramid module. The decoder network upsamples the feature maps to the original input resolution and fuses them with the corresponding feature maps from the backbone network, using a skip-connection mechanism similar to that used in U-Net [17, 18].

The final output of DeepLabV3+ is a pixel-wise segmentation map, where each pixel is assigned a label indicating the object or background class. The model is trained using a cross-entropy loss function, which measures the discrepancy between the predicted and ground-truth segmentation maps.

Overall, DeepLabV3+ has demonstrated state-of-the-art performance on several benchmark datasets for semantic segmentation. Its combination of deep CNNs, spatial

pyramid pooling, and decoder networks allows it to achieve highly accurate segmentation results, even in complex scenes with multiple objects and cluttered backgrounds.

3.4 Performance Evaluation

There are several commonly used measures for evaluating the performance of image segmentation algorithms. Some of the most widely used measures are considered in this research work and the necessary expressions are shown in Eqns. (2) to (5) [19, 20];:

$$\text{Jaccard} = \text{TP}/(\text{TP} + \text{FP} + \text{FN}) \quad (2)$$

$$\text{Dice} = (2 * \text{TP})/(2 * \text{TP} + \text{FP} + \text{FN}) \quad (3)$$

$$\text{Accuracy} = (\text{TP} + \text{TN})/(\text{TP} + \text{TN} + \text{FP} + \text{FN}) \quad (4)$$

$$\text{Precision} = \text{TP}/(\text{TP} + \text{FP}) \quad (5)$$

where TP is the number of true positives, TN is the number of true negatives, FP is the number of false positives, and FN is the number of false negatives.

4 Result and Discussions

This section of the research presents the experimental outcome achieved using the implemented pretrained schemes. The proposed scheme is trained using 800 images along with its GT and the merit of this scheme is verified using 200 testing images.

The experimental outcome of DeepLabV3+ is presented in Fig. 3. Figure 3(a) presents the outcome achieved during algorithm training and Figs. 3(b) and (c) shows the training and validation accuracy and loss for a chosen epoch of 120. This confirms that the proposed approach is efficient in achieving a better accuracy. After the training, the proposed scheme provides a binary segmentation result and it is depicted in Fig. 3(d). This picture also confirms that the segmented outcome of this scheme is better and it is similar to the GT.

The performance measures from these images need to be computed by executing a pixel wise comparison among the segmented section and the GT and the mean value of these measures are then considered for the validation. The sample results and the related GT obtained during the study is presented in Fig. 4.

The result of DeepLabV3+ is compared and confirmed against other segmentation procedures, like UNet, SegNet, VGG-UNet and VGG-SegNet available in the literature and the mean \pm standard deviation value is then considered for the assessment as shown in Table 1. This table confirms that, the proposed approach helps in getting a better result on the chosen database compared to other existing approaches.

The limitation of this research is, it implements a thresholding combined with the segmentation process. In the future, the SE can be replaced with other thresholding schemes existing in the literature. Further, the computational complexity of the proposed scheme can be reduced by considering the light-weight deep learning based segmentation schemes.

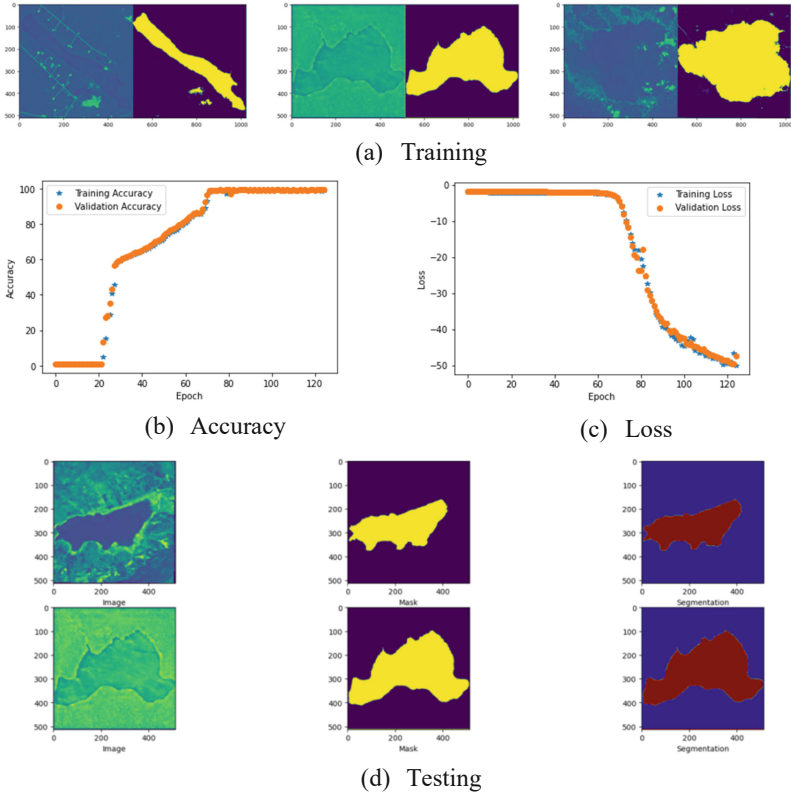


Fig. 3. Experimental outcome for DeepLabV3+

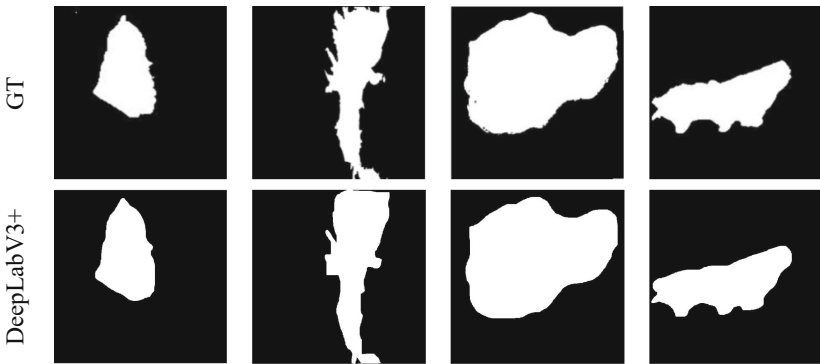


Fig. 4. Sample GT section and its related segmented waterbody region

Table 1. Computed performance values of the test images

Scheme	Jaccard (%)	Dice (%)	Accuracy (%)	Precision (%)
DeepLabV3+	89.28 ± 1.05	93.63 ± 0.85	97.38 ± 0.31	95.17 ± 0.22
UNet	87.17 ± 1.17	91.13 ± 0.82	96.06 ± 0.33	93.37 ± 0.41
SegNet	86.81 ± 1.08	90.28 ± 0.75	96.22 ± 0.28	93.09 ± 0.38
VGG-UNet	88.83 ± 1.15	92.53 ± 0.38	97.14 ± 0.27	94.85 ± 0.16
VGG-SegNet	88.39 ± 1.13	92.17 ± 0.52	96.74 ± 0.19	94.29 ± 0.23

5 Conclusion

This work implements a methodology by integrating the thresholding and de-segmentation procedure to detect waterbodies in SI. The experimental outcome confirms that the DeepLabV3+ has demonstrated its ability to accurately detect waterbodies, even in complex scenes where the waterbodies are partially or completely occluded. By leveraging the powerful features of DeepLabV3+ and fine-tuning it on waterbody detection datasets, the model can accurately detect waterbodies at different scales, shapes, and orientations. Additionally, DeepLabV3+ can effectively distinguish waterbodies from other similar-looking features such as shadows, rocks, and vegetation, making it a robust tool for waterbodies detection. Overall, DeepLabV3+ based waterbodies detection has the potential to provide valuable insights into water resources management, environmental monitoring, and disaster response planning. However, like any other computer vision model, it has its limitations and requires careful evaluation and validation before deployment in real-world applications.

References

1. Kadhim, I.J., Premaratne, P.: A novel deep learning framework for water body segmentation from satellite images. *Arabian J. Sci. Eng.* 1–12 (2023)
2. Rambhad, A., Singh, D.P., Choudhary, J.: Detection of flood events from satellite images using deep learning. In: *Intelligent Data Engineering and Analytics: Proceedings of the 10th International Conference on Frontiers in Intelligent Computing: Theory and Applications (FICTA 2022)*, pp. 259–268. Singapore: Springer Nature Singapore (2023)
3. Daniel, J., Rose, J.T., Vinnarasi, F., Rajinikanth, V.: VGG-UNet/VGG-SegNet supported automatic segmentation of endoplasmic reticulum network in fluorescence microscopy images. *Scanning* 2022 (2022)
4. Rajinikanth, V., Kadry, S., Damaševičius, R., Sankaran, D., Mohammed, M.A., Chander, S.: Skin melanoma segmentation using VGG-UNet with Adam/SGD optimizer: a study. In: *2022 Third International Conference on Intelligent Computing Instrumentation and Control Technologies (ICICT)*, pp. 982–986. IEEE (2022)
5. Zhang, Z., Meng, L., Ji, S., Huaifen, Y., Nie, C.: Rich CNN features for water-body segmentation from very high resolution aerial and satellite imagery. *Remote Sensing* **13**(10), 1912 (2021)
6. Guo, Z., Lin, W., Huang, Y., Guo, Z., Zhao, J., Li, N.: Water-Body Segmentation for SAR Images: Past, Current, and Future. *Remote Sensing* **14**(7), 1752 (2022)

7. Yuan, K., Zhuang, X., Schaefer, G., Feng, J., Guan, L., Fang, H.: Deep-learning-based multispectral satellite image segmentation for water body detection. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **14**, 7422–7434 (2021)
8. Rana, H., Sivakumar Babu, G.L.: Object-oriented approach for landslide mapping using wavelet transform coupled with machine learning: a case study of Western Ghats, India. *Indian Geotechnical Journal* **52**(3), 691–706 (2022)
9. Aalan Babu, A., Mary Anita Rajam, V.: Water-body segmentation from satellite images using Kapur's entropy-based thresholding method. *Computational Intelligence* **36**(3), 1242–1260 (2020)
10. <https://www.kaggle.com/datasets/franciscoescobar/satellite-images-of-water-bodies?select=Water+Bodies+Dataset>
11. Kadry, S., Rajinikanth, V., Koo, J., Kang, B.-G.: Image multi-level-thresholding with Mayfly optimization. *Int. J. Elect. Comp. Eng.* **11**(6), 2088–8708 (2021)
12. Rajinikanth, V., Palani Thanaraj, K., Satapathy, S.C., Fernandes, S.L., Dey, N.: Shannon's entropy and watershed algorithm based technique to inspect ischemic stroke wound. In: *Smart Intelligent Computing and Applications: Proceedings of the Second International Conference on SCI 2018*, Vol. 2, pp. 23–31. Springer Singapore (2019)
13. Rajinikanth, V., Satapathy, S.C., Dey, N., Vijayarajan, R.: DWT-PCA image fusion technique to improve segmentation accuracy in brain tumor analysis. In: *Microelectronics, Electromagnetics and Telecommunications: Proceedings of ICMEET 2017*, pp. 453–462. Springer Singapore (2018)
14. Zervoudakis, K., Tsafarakis, S.: A mayfly optimization algorithm. *Comput. Ind. Eng.* **145**, 106559 (2020)
15. Bhattacharyya, T., et al.: Mayfly in harmony: a new hybrid meta-heuristic feature selection algorithm. *IEEE Access* **8**, 195929–195945 (2020)
16. Zhang, D., Ding, Y., Chen, P., Zhang, X., Pan, Z., Liang, D.: Automatic extraction of wheat lodging area based on transfer learning method and deeplabv3+ network. *Comput. Electron. Agric.* **179**, 105845 (2020)
17. Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F., Adam, H.: Encoder-decoder with atrous separable convolution for semantic image segmentation. In: *Proceedings of the European conference on computer vision (ECCV)*, pp. 801–818 (2018)
18. Peng, H., Zhong, J., Liu, H., Li, J., Yao, M., Zhang, X.: ResDense-focal-DeepLabV3+ enabled litchi branch semantic segmentation for robotic harvesting. *Comput. Electron. Agric.* **206**, 107691 (2023)
19. Rajinikanth, V., Durai Raj Vincent, P.M., Srinivasan, K., Ananth Prabhu, G., Chang, C.-Y.: A framework to distinguish healthy/cancer renal CT images using the fused deep features. *Frontiers in Public Health* **11** (2023)
20. Manic, K.S., Rajinikanth, V., Al-Bimani, A.S., Taniar, D., Kadry, S.: Framework to Detect Schizophrenia in Brain MRI Slices with Mayfly Algorithm-Selected Deep and Handcrafted Features. *Sensors* **23**(1), 280 (2022)