ISSN: 0976-9102 (ONLINE)
DOI: 10.21917/ijivp.2025.0500

AUTOMATIC AND FAST RIVER WATER DETECTION SYSTEM BASED DCONVLU-NET FROM REMOTE SENSING IMAGERY

Miral J Patel[1], Hasmukh P Koringa[2], Jay Pandya[3] and Darshan Tank[4]
[1,2,3] Department of Electronics and Communication, Government Engineering College, Rajkot, India
[4] Department of Information Technology, Government Polytechnic, Rajkot, India

Abstract

River segmentation from Remote Sensing Imagery (RSI) has significant research value and practical applications for monitoring river changes, comprehending patterns in river water levels, flood detection, agricultural planning, and environmental monitoring. Therefore, monitoring river areas and water bodies is essential. This paper proposes the deep-learning approaches based on a decreased Convolutional Layer U-Net based (dConvLU-Net) method to perform an efficient segmentation of river and land from RSI containing inference of non-river information such as bridges, shadows, and roads. The results of the experiments show how well these models work compared with other semantic segmentation models in many aspects of river water segmentation. The FCN-based method takes less execution time and the least computational cost but the mean Pixel Accuracy(mPA) and Mean Intersect over Union (mIOU) are also less. U-Net performs better mPA and mIoU despite their increased computational costs and execution time. However, the dConvLU-Net method performs most effectively regarding execution time, computational cost, mPA, and mIoU. The results of the proposed dConvLU-Net method show that the river segmentation from RSI is fast and accurate.

Keywords:

Deep Learning, Segmentation, RSI, U-Net

1. INTRODUCTION

Water bodies especially rivers are vital to the existence of all life forms on Earth. In the past, the monitoring relied essentially on manpower in surveying individual areas. However, there are limitations associated with such surveys, e.g., the tremendous amount of time and labor involved in expeditions. Presently, there has been accelerated development in remote sensing and artificial intelligence (AI) technology, particularly for water source monitoring and change detection in different areas. In recent years, deep learning methods for semantic segmentation have been the preferred choice given their high accuracy and ease of use. Convolutional neural network-based deep learning techniques are helpful for feature extraction [1], semantic segmentation [2] [3], and image classification [4] [5]. The convolutional neural network-based method does not require human intervention for modeling because of its automatic feature learning capacity.

Earlier days water segmentation from RSI was used image processing techniques. Using Gabor Filtering and Path Opening method. It is non automatic approach to segment a river from the RSI [6]. A Fully Convolutional Network (FCN) [7] handles the problem of image segmentation at the semantic level, classifies images at the pixel level, and transforms the network's last three layers into a 1×1 convolutional kernel. To get additional spatial information construct the U-Net network based on FCN, use de-convolution as the up-sampling structure, and accomplish feature information fusion through the skip connections [8]. The proposed method of water segmentation is based on transfer learning and water level measurement. ADE20k and COCO-stuff datasets were used for water segmentation using the semantic segmentation algorithm of FCNs and DeepLab with ResNet50 and ResNet101 encoder respectively [9]. Applied method on the Water Segmentation Open Collection (WSOC) dataset with a combination of three different backbones (VGG16, ResNet50, and MobilNet) and four various deep learning methods. SegNet and UNet both perform well when backbone as ResNet. SegNet required less time to segment water from the background image so it was used for emergencies of flood detection and monitoring[10]. An exhaustive comparison of U-Net, PSPNet, DeeplabV3+, PAN, and LinkNet backboned with predefined ResNet50 and SAM in the river water segmentation task from remote sensing images. The experimental results were conducted on benchmark river water segmentation and the LuFI-RiverSnap.v1 dataset. The U-Net(ResNet50) was more accurate on average than the other tested models in river water segmentation[11]. An efficient extraction method proposed based on a composite attention mechanism for river segmentation [12]. The suggested method has certain advantages as the training process can converge quickly and all indexes are superior but it fails to detect river boundaries, small buildings along the bank, and bridges, it is still unable to predict correctly. SegNet based method automatically segment river water in imagery acquired by RGB sensors [13]. It measures pixel accuracy and IoU to segment the water from input image DeepRivWidth method proposed for the Deep learning-based semantic segmentation approach for river identification and width measurement in SAR images of Coastal Karnataka[14]. The proposed method detects very narrow rivers accurately. The proposed method detect river channel changes from the remote sensing images [15]. Ensemble learning approach that leverages the unique representations learned by each backbone, resulting in more robust and accurate segmentation [16]. It was analyzed by IoU values, confusion matrices, and sample inference masks. Vanilla U-Net and Transfer U-Net method suggested for river water detection by [17]. Extraction of water from high-resolution remote RSI based on a deep semantic segmentation network suggested by [18]. Lots of researcher suggested the Efficient river identification from high resolution images used for various applications such as unmanned surface vehicles suggested [19], riverine litter [20] . An automatic planning of power transmission lines needs to recognize the feature such as river, road, and buildings information from RSI [21]. Due to continuously improvement in Remote sensing satellites are gradually improving their image resolution. Therefore feature extraction with use of a deep learning and graphics processing units makes easy to identified object from RSI.

2. MATERIALS AND METHOD

2.1 PROPOSED METHODOLOGY

Many researchers propose methods to identify the river water bodies from the RSI. However, the river segmentation of RSI images is challenging due to identical objects present such as roads, occlusion, bridges, and river banks. To overcome these problems, we propose a river segmentation method called dConvLU-Net (Decrease Convolutional Layers), which is built on a U-Net. Originally U-Net was designed for medical diagnosis purposes that required to segment of very minute objects from the patient's image. However, here we applied this method for detecting large objects like rivers. So here we decrease the number of convolution layers. The architecture of the proposed dConvLU-Net is shown in Fig.1. The proposed method has only 13 convolution layers than the original U-Net. As a result, the model's training parameters are decreased. Consequently, the model takes less time to train than the original U-Net and requires fewer computation resources. However, it can identify the river bodies from RSI with around the same segmentation efficiency.

2.2 PERFORMANCE METRICS MEASUREMENT

For semantic segmentation efficacy, performance measurements such as accuracy, precision, recall, IOU, and F1 score are typically employed. They were based on the confusion matrix of the test dataset. They defined as:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$\text{Recall} = \frac{TP}{TP + FN}, \quad (2)$$

$$\text{Precision} = \frac{TP}{TP + FP}, \quad (3)$$

$$\text{F1score} = 2 * \text{Recall} \frac{\text{precision}}{\text{precision} + \text{Recall}}, \quad (4)$$

$$\text{IOU} = \frac{TP}{TP + FP + FN}, \quad (5)$$

where True Positive (TP) represents the number of river pixels classified as correctly, True Negative (TN) refers to the number of background pixels other than river pixels classified as correctly, False Positive (FP) shows the number of the incorrectly classified river pixels, and False Negative (FN) denoted that number of the background other than river pixels are classified wrongly. Accuracy displays the model's overall performance across all classes. The number of correctly identified positive samples divided by the total number of positive samples in the test set is known as recall. The percentage of correctly predicted positive samples among all expected positive samples is known as an image's precision. The precision and recall harmonic average values are represented by the F1 score. The IOU is a representation of the predicted and original image overlap ratio.

2.3 DATASET: RIWA_V2 [22]

This is a powerful and adaptable dataset for river scene analysis and segmentation, with 1128 images taken with UAVs and cell phones [22]. Because the rivers in the dataset are diverse in hue, it can be difficult for models to distinguish between water and background. Total 866 images are used to train the River Water Bodies Segmentation model. 70% images are used for training phase and 15% of images are used for validation phase and 15% images are used testing phase of the River Segmentation Model.

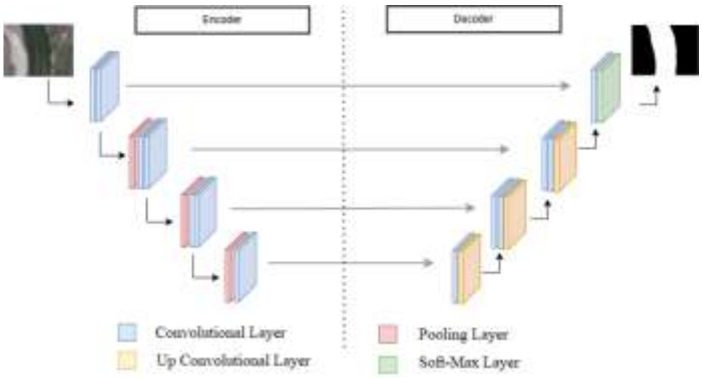

Fig.1. dConvLU-Net: U-Net Based Architecture

3. RESULT AND ANALYSIS

3.1 EXPERIMENT SETUP

The deep learning library in PyTorch is used to implement the River water segmentation based on the U-Net model in Python. The Google Colab is used to train the model. Google Colab is a GPU-accelerated With an NVIDIA T4 Tesla GPU, 12 GB of RAM, CUDA version 12.2, and an Intel Xeon CPU running at 2.00 GHz with two threads and one core. The input dataset split into train, validation and test dataset. The data augmentation was applied on the training set and validation set of the dataset. Data augmentation used techniques of the horizontal flip, vertical flip, and rotate with 450.

Algorithm 1 DL Model for Training Phase of River Segmentation

Input : $i_{train} = \{(I_i, G_i)\}_{i=0}^{N_{train}-1}$, $i_{val} = \{(I_i, G_i)\}_{i=0}^{N_{val}-1}$

Learning Rate £, Batch Size β, Number of Epochs N_{epochs}

Output: Model f_m, weight Parameters $\varnothing_m$

Initialization: Initialize Model Weight Parameter $\varnothing_m$

Input Pre-Processing :
for each sample of the train set **do**
$(I_i^A, G_i^A) \leftarrow Augment\, i_{train}$
for each sample of the validation set **do**
$(I_i^A, G_i^A) \leftarrow Augment\, i_{val}$
Training Phase:
for epochs 1 **to** N_{epochs} **do**
for each **minibatch** $(I_{i:i+\beta-1}^A, G_{i:i+\beta-1}^A) \in i_{train}$
$\hat{G}_{i:i+\beta-1}^A \leftarrow f_m(I_{i:i+\beta-1}^A; \varnothing_m)$
$Ł_{BCE(i:i+\beta-1)} \leftarrow BCE(\hat{G}_{i:i+\beta-1}^A, G_{i:i+\beta-1}^A)$

$$\varnothing_m \leftarrow Adam(\theta_m, \nabla_{\varnothing_m} \frac{1}{\beta}\sum_{j=i}^{i+\beta-1} Ł_{BCE_j}, \lambda)$$

Validation Phase:

for epochs 1 **to** N_{epochs} **do**

for each **minibatch** $(I^A_{i:i+\beta-1}, G^A_{i:i+\beta-1}) \in i_{\text{val}}$

$\hat{G}^A_{i:i+\beta-1} \leftarrow f_m(I^A_{i:i+\beta-1}; \varnothing_m)$

$Ł_{BCE(i:i+\beta-1)} \leftarrow BCE(\hat{G}^A_{i:i+\beta-1}, G^A_{i:i+\beta-1})$

Calculate Evaluations Matrices

for i← 0 to N_{epochs}-1 **do**

mIo←avg(IoU)

mPA←avg(PA)

Execution Time, Total Parameters

end for

The proposed model used Algorithm 1 during the training phase and Algorithm 2 during the testing phase. The Learning Rate £ of the model is initialized by the 1x10-4, Batch Size β of 4, Hidden Layers of 64, and image size of 512x512 pixels. Drop-out layer probabilities of p1 and p2 are set to 0.25 and 0.5 respectively. The proposed model used a Binary Cross Entropy (BCE) loss model and Adam optimizer to optimize the weight parameter of the dConvLU-Net model. During the training phase of the model (FCN,U-Net and dConvLU-Net) calculate the mPA, mIoU, Loss, and Execution time of the model. During the testing phase of the model FCN, U-Net, and dConvLU-Net) calculate the recall, precision, and F1 score of the model.

Algorithm 2 DL Model for Testing Phase of River Segmentation

Input: $i_{\text{test}} = \{(I_i, G_i)\}_{i=0}^{N_{\text{test}}-1}$, Trained River Semantic Segmentation Model f_m, weight Parameters $\varnothing_m$, $i_{\text{val}} = \{(I_i, G_i)\}_{i=0}^{N_{\text{val}}-1}$, Learning Rate £, Batch Size β, Number of Epochs N_{epochs}

Output: Predicted Output of test dataset $\hat{G}_{\text{test}}$, Average Evaluation Matrix

Initialization: Initialize Model Weight Parameter $\varnothing_m$

Metrics_List← [ ]

Calculate Evaluations matrices

for i←0 to N_{test}-1 do

$\hat{G}^A_i \leftarrow f_m(I^A_i; \varnothing_m)$; /* Trained model*/

Metrics_List.append(Evaluated_Metrics $(\hat{G}^P_i, G^P_i)$) ;

$\hat{G}_{\text{test}} \leftarrow \hat{G}^P_i$

Calculate Average Evaluation Metrics

Mean_Evaluation ←Mean(Metrics_List)

Return predicted test output $\hat{G}_{\text{test}}$, Average Evaluations metrics of Recall, Precision and F1 score

3.2 RIVER SEGMENTATION TRAINING PROCESS

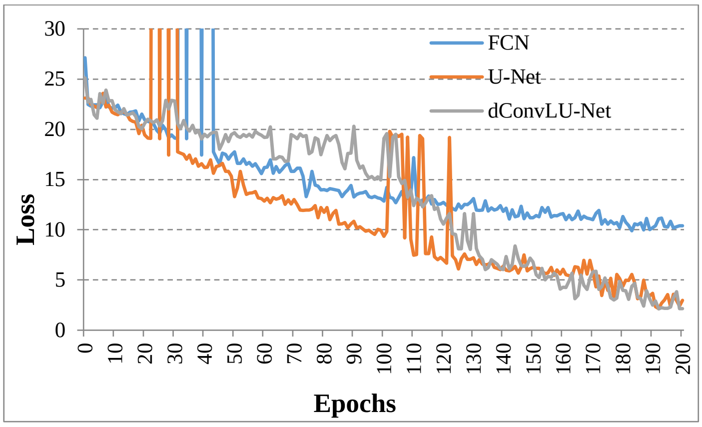

Fig.2. Loss Plot for FCN, U-Net, and dConvLU-Net

The parameters for the three semantic segmentation algorithms (FCN, U-Net, and U-Net Based dConvLU-Net) are learned on the RiWa V2 dataset. Training of the model is done according to the algorithm 1. The identical training, validation, and testing dataset is used to train each of the three models independently. Training for the FCN, U-Net, and dConvLU-Net took place across 200 epochs. The loss function was applied for binary cross entropy since this is a binary class segmentation problem. Adam's optimizer was used with a constant learning rate of 1×10^{-4}. The weights were initialized by Xavier Initialization [23]. The Fig.2 shows the loss plot for the FCN, U-Net, and dConvLU-Net. The Average Pixel Accuracy measurement is done for each model. Fig.3 shows the plot for all three models of average pixel accuracy against each epoch. It shows the average accuracy of the FCN model is very low compared to the other two U-Net-based models. The accuracy of the U-Net and dConvLU-Net is approximately the same but the training time required by U-Net is more compared to the dConvLU-Net. From Fig.4, we can conclude that the IoU score of the FCN is poor compared with the other two deep learning based methods.

In Table.1, the overall accuracy of the various approaches throughout the training phase is listed along with the IoU score. Additionally, it displays the FCN, U-Net, and dConvLU-Net model execution times for river identification. For 200 epochs, the FCN-based model takes 4 hours and 22 minutes to execute, whereas U-Net takes 5 hours and 48 minutes. In contrast to the original U-Net, the dConvLU-Net model contains fewer convolutional layers; hence it took 4 hours and 28 minutes to train the model. The proposed dConvLU-Net has IoU score of 91.33 % whereas U-Net has IoU of 90.76%. Because a leaky Relu activation function was utilized in place of Relu, the accuracy and IoU of the suggested DconvLU-Net improved throughout the training phase. The Leaky Relu activation function tuned the hyper parameters more precisely as it provide the value of negative input [24].

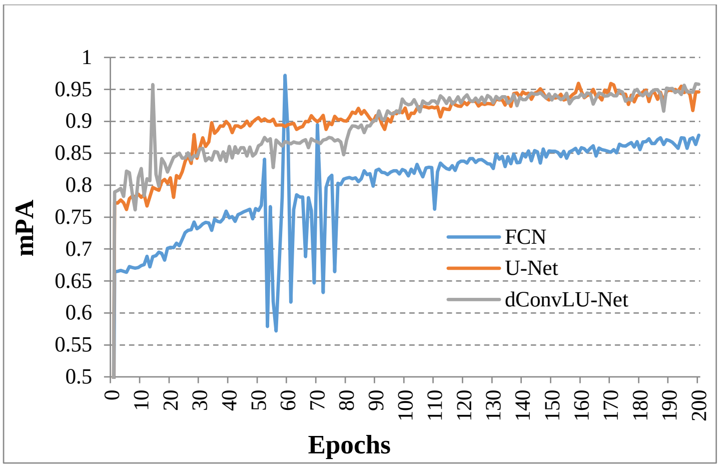

Fig.3. Average Accuracy Plot for FCN, U-Net and dConvLU-Net

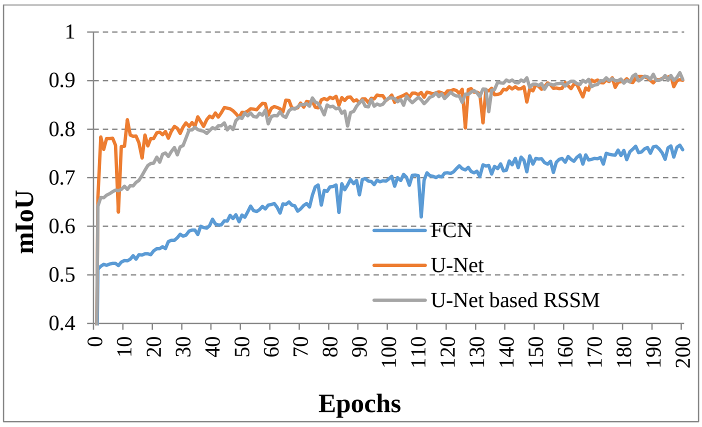

Fig.4. Average IoU Plot for FCN, U-Net, and dConvLU-Net

Table.1. Training phase parameters

Model	Execution Time	Accuracy (%)	IoU (%)
FCN	4 hr 22 min	87.82	76.54
U-Net	5 hr 48 min	95.11	90.76
dConvLU-Net	4 hr 28 min	95.68	91.33

3.3 TESTING AND ANALYSIS OF RIVER WATER BODIES IDENTIFICATION

Table.2. Testing Phase Parameters

Model	Precision (%)	Recall (%)	F1 score (%)
FCN	81	78	79.47
U-Net	86	92	88.89
dConvLU-Net	88	91	89.47

A comprehensive summary of the performance analysis of the FCN, U-Net, and dConvLU-Net during the testing phase of the models has been presented in Table 2. The FCN model has less computation time but the F1 score among the three was 79.47 % was least among three FCN, U-Net, and dConvLU-Net methods. . However, U-Net has the highest Recall of 92 % to train the river identification model among three deep learning-based methods. The dConvLU-Net model has moderate Recall of 91 % and precision of the 88% that made the highest F1 score, as well as the model, needs less computation time during the training phase of the river water bodies identification as displayed in Table 1. The dConvLU-Net model is the fastest, automatic and most suitable for river identification from the aerial images as shown in Fig.5. For all test samples, the FCN image was identified images poorly as lots of the pixels are incorrectly identified. Most of the test images both U-Net and dConvLU-Net identified approximate similar river identification as shown in Fig.5. However, the U-Net-based method failed to identify the complex image as sharp curves of the river were present whereas the dConvLU-Net method did it as tuning of hyper-parameters was well done for the complex images too.

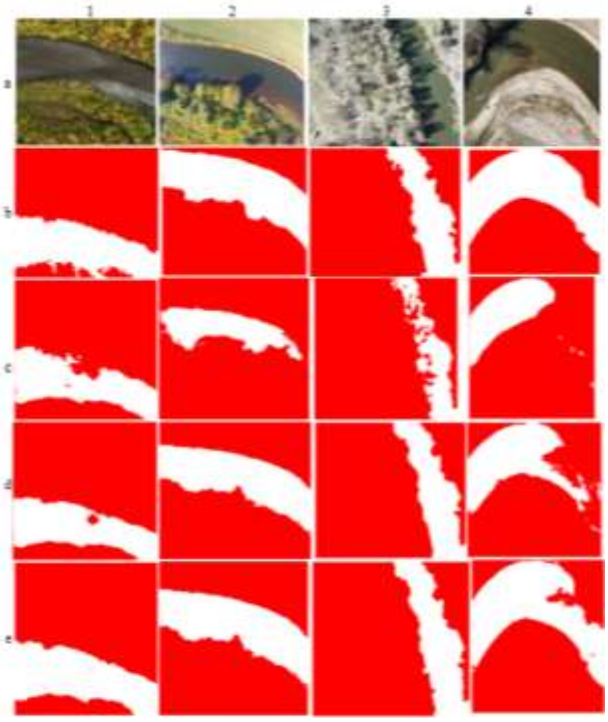

Fig.5. Some Test Output of River Identification Result (a) Original Image (b) Ground Truth (c) FCN (d) U-Net (e) dConvLU-Net.

4. CONCLUSION

This study provided a thorough comparison of the proposed U-Net-based model dConvLU-Net, FCN, and U-Net in the river water segmentation. The experimental findings were obtained using the benchmark RiWa_V2 dataset for river water segmentation. These experiments provide the models' effectiveness and versatility on the testing dataset. This insightful information is used for further developments in river water segmentation studies and applications. Based on the experimental results, it is possible to segment river water from RSI with significant differences in water color, illumination, sky, structure reflection on the surface water, and sharp curves using a variety of models. Compared to the other two approaches, the FCN took less computing time to train the model, but its F1 score was the lowest. In terms of river water segmentation, the U-Net was generally more accurate than the other models that were tested; nevertheless, its computation time was slower. The U-Net had more Recall value than the other tested models in river water segmentation; however, it was slower in computing. The dConvLU-Net achieves the highest F1 score. It was the most efficient model regarding computation time. From Fig.5 it was concluded the dConvLU-Net method provides river water identification from remote sensing images in complex scenarios such as sharp curves of rivers are present in the input image.

REFERENCES

[1] M. Xia, X. Zhang, W. Liu, L. Weng and Y. Xu, "Multi-Stage Feature Constraints Learning for Age Estimation", *IEEE Transactions on Information Forensics and Security*, Vol. 15, No. 1, pp. 2417-2428, 2020.

[2] Patel, Miral Jerambhai and M. Ashish Kothari, "Deep Learning-Enabled Road Segmentation and Edge-Centerline Extraction from High-Resolution Remote Sensing Images", *International Journal of Image and Graphics*, Vol. 23, No. 6, pp. 1-7, 2023.

[3] Hasmukh Koringa and Miral Patel, "Automated, Quick and Precise Building Extraction From Aerial Images using Ll-Unet Model", *Radio Electronics and Computer System*, Vol. 2024, No. 2 pp. 41-51, 2024.

[4] J. Liang, Y. Deng and D. Zeng, "A Deep Neural Network Combined CNN and GCN for Remote Sensing Scene Classification", *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, Vol. 13, No. 99, pp. 1-7, 2020.

[5] P. Deng, H. Huang and K. Xu, "A Deep Neural Network Combined with Context Features for Remote Sensing Scene Classification", *IEEE Geoscience and Remote Sensing Letters*, No. 99, pp. 1-5, 2020.

[6] K. Yang, M. Li, Y. Liu, L. Cheng, Q. Huang and Y. Chen, "River Detection in Remotely Sensed Imagery using Gabor Filtering and Path Opening", *Remote Sensing*, Vol. 7, No. 7, pp. 8779-8802, 2015.

[7] J. Long, E. Shelhamer and T. Darrell, "Fully Convolutional Networks for Semantic Segmentation", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 39, No. 4, pp. 640-651, 2015.

[8] O. Ronneberger, P. Fischer and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation", *Computer Vision and Pattern Recognition*, Vol. 67, No. 1, pp. 1-7, 2015.

[9] R. Vandaele, S.L. Dance and V. Ojha, "Automated Water Segmentation and River Level Detection on Camera Images using Transfer Learning", *Proceedings of International Conference on Pattern Recognition*, pp. 232-245, 2020.

[10] M. Zaffaroni and C. Rossi, "Water Segmentation with Deep Learning Models for Flood Detection and Monitoring", *International Journal of Information Systems for Crisis Response and Management*, Vol. 12, No. 1, pp. 24-27, 2020.

[11] A. Moghimi, M. Welzel, T. Celik and T. Schlurmann, "A Comparative Performance Analysis of Popular Deep Learning Models and Segment Anything Model for River Water Segmentation in Close-Range Remote Sensing Imagery", *IEEE Access*, Vol. 12, pp. 1-6, 2024.

[12] Z. Fan, J. Hou, Q. Zang, Y. Chen and F. Yan, "River Segmentation of Remote Sensing Images based on Composite Attention Network", *Complexity*, Vol. 45, No. 1, pp. 1-6, 2022.

[13] T.S. Akiyama, J. Marcato Junior, W.N. Gonçalves, P.O. Bressan, A. Eltner, F. Binder and T. Singer, "Deep Learning Applied to Water Segmentation", *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. 43, pp. 1189-1193, 2020.

[14] U. Verma, A. Chauhan, M.M. Manohara Pai and R. Pai, "DeepRivWidth: Deep Learning based Semantic Segmentation Approach for River Identification and Width Measurement in SAR Images of Coastal Karnataka", *Computers and Geosciences*, Vol. 154, pp. 1-7, 2021.

[15] R. Vandaele, S.L. Dance and V. Ojha, "Deep Learning for Automated River-Level Monitoring through River-Camera Images: an Approach based on Water Segmentation and Transfer Learning", *Hydrology and Earth System Sciences*, Vol. 25, No. 8, pp. 4435-4453, 2021.

[16] I. Dimitrovski, V. Spasev, S. Loshkovska and I. Kitanovski, "U-Net Ensemble for Enhanced Semantic Segmentation in Remote Sensing Imagery", *Remote Sensing*, Vol. 16, No. 12, pp. 1-6, 2024.

[17] M.M. Pai, V. Mehrotra, S. Aiyar, U. Verma and R.M. Pai, "Automatic Segmentation of River and Land in SAR Images: A Deep Learning Approach", *Proceedings of International Conference on Artificial Intelligence and Knowledge Engineering*, pp. 15-20, 2019.

[18] D. Sun, G. Gao, L. Huang, Y. Liu and D. Liu, "Extraction of Water Bodies from High-Resolution Remote Sensing Imagery based on a Deep Semantic Segmentation Network", *Scientific Reports*, Vol. 14, No. 1, pp. 1-6, 2024.

[19] J. Zhang, J. Gao, J. Liang, Y. Wu, B. Li, Y. Zhai and X. Li, "Efficient Water Segmentation with Transformer and Knowledge Distillation for USVs", *Journal of Marine Science and Engineering*, Vol. 11, No. 5, pp. 1-7, 2023.

[20] A. Mohsen, T. Kiss and F. Kovács, "Machine Learning-based Detection and Mapping of Riverine Litter Utilizing Sentinel-2 Imagery", *Environmental Science and Pollution Research*, Vol. 30, No. 25, pp. 67742-67757, 2023.

[21] G. Nan, H. Li, H. Du, Z. Liu, M. Wang and S. Xu, "A Semantic Segmentation Method based on AS-Unet++ for Power Remote Sensing of Images", *Sensors*, Vol. 24, No. 1, pp. 1-7, 2024.

[22] X. Blanch, F. Wagner and A. Eltner, "RIWA Dataset", Available at https://www.kaggle.com/dsv/4289421, Accessed in 2022.

[23] X. Glorot and Y. Bengio, "Understanding the Difficulty of Training Deep Feedforward Neural Networks", *Proceedings of International Conference on Artificial Intelligence and Statistics*, pp. 249-256, 2010.

[24] J. Xu, Z. Li, B. Du, M. Zhang and J. Liu, "Reluplex Made More Practical: Leaky ReLU", *Proceedings of International Symposium on Computers and Communications*, pp. 1-7, 2020.