Computers & Geosciences 154 (2021) 104805

Contents lists available at ScienceDirect

Computers and Geosciences

journal homepage: www.elsevier.com/locate/cageo

# DeepRivWidth : Deep learning based semantic segmentation approach for river identification and width measurement in SAR images of Coastal Karnataka

Ujjwal Verma [a], Arjun Chauhan [a], Manohara Pai M.M. [b,*], Radhika Pai [b,*]

[a] Department of Electronics & Communication Engineering, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, India
[b] Department of Information & Communication Engineering, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, India

ARTICLE INFO

*Keywords:*
Semantic segmentation
Synthetic Aperture Radar
River width measurement
Convolutional neural networks

ABSTRACT

River width is an essential parameter for studying the river's hydrological process and has been widely used to estimate the river discharge. The existing approaches to measuring river width are based on remotely sensed imagery such as MODIS, Landsat to identify the river, and then estimate the river width. In this work, an alternate approach for river width estimation is proposed using the under-explored modality Synthetic Aperture Radar (SAR) images. SAR, unlike the traditional electro-optical sensors, can penetrate the clouds and can be used to collect the data in all weather conditions and even during the night. In this work, the river identification process is manifested as a binary semantic segmentation task in SAR images. For this, two state of the art deep learning algorithms (U-Net, DeepLabV3+) are utilized for river identification and subsequent width measurement. The proposed approach (DeepRivWidth) is used to estimate the width of the river of the Mangalore–Udupi region of Coastal Karnataka (India). These rivers originate or pass through Western Ghats (UNESCO world heritage site), and the proposed river width measurement approach could provide critical input for ecologists besides assisting efficient water management of the region. The estimated width is compared with the manually measured width, and significant improvement in the accuracy was obtained compared to existing river width measurement approaches. Besides, the performance evaluation of semantic segmentation approaches for river identification on a publicly available dataset provides valuable insights into segmenting rivers in SAR images.

## 1. Introduction

River width is one of the crucial river parameters for studying the river's hydrological, biological, geological, and chemical processes (Ling et al., 2019). Besides, river width provides useful information for measuring carbon dioxide movement between rivers and the atmosphere. In addition, flow discharge can also be estimated from river width information (Ling et al., 2019; Gleason and Smith, 2014). The study of river width information can provide valuable insights into the changes in river flow and shape. It could also be used to study extreme climatic events such as floods and droughts. The temporal river width information could help identify soil erosion along the river banks over a certain period and thus assist government agencies in taking preventive measures at these locations.

Traditionally, the river width is measured in the field, which is a time-consuming and challenging task in rugged terrain. However, recent advances in remote sensing technologies have allowed the measurement of these quantities from space. Remote sensing technologies provide a cost-effective alternative path to measure these quantities for a more substantial area as compared to the traditional field measurement. For instance, RivWidth is one of the tools developed for measuring river width (Pavelsky and Smith, 2008). This tool uses 250-m Moderate Resolution Imaging Spectroradiometer (MODIS) data and U.S. Geological Survey National Land Cover Dataset as the input for identifying Lena and Ohio rivers, respectively. Subsequently, the centerline of the river is estimated using edge detection techniques. Finally, the river width is measured along the orthogonal direction at each centerline pixel. In another study, a Google Earth Engine based river width measurement tool is presented in Yang et al. (2020). The tool presented in Yang et al. (2020) utilizes spectral indices for river identification. Besides, Global River Widths from Landsat (GRWL) dataset is also utilized to further refine the river identification results. Subsequently, the center line of the river is computed using distance transform and skeletonization. Finally, the width of the river is measured along the

* Corresponding authors.
*E-mail addresses:* mmm.pai@manipal.edu (Manohara Pai M.M.), radhika.pai@manipal.edu (R. Pai).

https://doi.org/10.1016/j.cageo.2021.104805
Received 16 September 2020; Received in revised form 8 February 2021; Accepted 28 April 2021
Available online 12 May 2021

direction orthogonal to the centerline. In a separate study (Ling et al., 2019), the author proposes to use coarse resolution remotely sensed imagery (MODIS, Landsat) for river width estimation. This approach focuses on accurate river identification by utilizing a deep learning-based super-resolution mapping to generate a fine resolution river map.

As discussed in Yang et al. (2020), the accuracy of river width measurement heavily depends upon the robustness of the method used for the identification of rivers. Thresholding is used as the existing approach for river identification, where the threshold is determined experimentally on a subset of data (Pavelsky and Smith, 2008; Yang et al., 2020). Besides, these methods rely on MODIS, Landsat images, which are affected by cloud presence (Li et al., 2014; Mateo-García et al., 2017). In this work, an alternative approach for river identification is proposed by utilizing under-explored modality Synthetic Aperture Radar (SAR) images. Unlike the traditional electro-optical sensors, SAR can penetrate the clouds and can be used to collect the data in all weather conditions and even during the night. Note that river width measurement algorithms proposed in Yang et al. (2020) utilizes a separate post-processing step to identify the presence of obstructions such as clouds, clouds shadow, snow, ice, and other natural factors. In comparison, the use of SAR image in measuring river width eliminates the need for such post-processing steps.

The authors in Klemenjak et al. (2012) proposed a mathematical morphological (path opening, closing) based approach for identifying rivers from TerraSAR-X data. However, determining the parameters of path opening and closing (size, shape of structuring element) is a challenging task and can severely affect the results obtained. In a separate study, a watershed based segmentation approach is presented in Ciecholewski (2017) for river identification in Polarimetric SAR. However, this method requires a manual identification of the seed region for an accurate river identification. Texture analysis based approach has also been utilized to identify the homogeneous regions of rivers (Sghaier et al., 2016). However, the texture descriptor needs to be compared with a predetermined threshold to identify the homogeneous region. Indeed, identifying the appropriate thresholding parameter is challenging and the accuracy of the method significantly depends on the correct determination of this parameter. In contrast, the proposed method utilizes an end-to-end CNN architecture for river identification thus eliminating the need to design hand crafted features, and post-processing the obtained results with the morphological operators.

Further, this work formulates river identification as a semantic segmentation problem. Segmentation is a widely studied topic in computer vision and has been used for various applications such as biomedical image analysis (Ronneberger et al., 2015), autonomous driving (Cordts et al., 2016), UAV aerial video analysis (Girisha et al., 2019), and precision agriculture (Verma et al., 2015, 2014). Semantic segmentation refers to the process of assigning a class label to each individual pixel in the image. However, the use of semantic segmentation on satellite images has been limited to the identification of oil spill, ships, buildings, roads, and similar objects. Wurm et al. (2019), Li et al. (2019), Zhou et al. (2018) and Song et al. (2020). To the best of our knowledge, there are limited works on the semantic segmentation of rivers in SAR images (Pai et al., 2019). Further, the river width has never been measured from SAR images. This work demonstrates that the river width can be measured precisely in SAR images. Besides, the process of width measurement has been simplified in this work by applying a novel approach for river identification using semantic segmentation on SAR images. The use of SAR images in river width estimation ensures accurate measurements in all weather conditions.

This work proposes an end-to-end system for river width estimation from SAR images (Fig. 1). In this work, the SAR images of the Mangalore–Udupi region of Southern India are first segmented using state-of-the-art semantic segmentation algorithms. This segmentation map is subsequently utilized to measure the width of rivers of the same area. The SAR images, along with its labels, is publicly available at https://github.com/ArjunChauhan0910/DeepRivWidth.

The main contributions of our work are:

- A river width measurement approach is presented from SAR images. To the best of our knowledge, there is no existing work on measuring river width from SAR images. Besides, the performance of two state of the art semantic segmentation algorithms are studied for river identification and the challenges encountered in river identification are highlighted in this work.
- Further, the dataset used in this study will be publicly available. We believe that the performance evaluation of two semantic segmentation algorithms on this dataset would serve as a baseline and further drive research for river segmentation and width measurement in SAR images. To the best of our knowledge, this will be the first publicly available dataset for river identification and width measurement from SAR images.
- The study area of this work are the rivers which originate or pass through the ecologically sensitive Western Ghats (Southern India), a UNESCO World Heritage site. To the best of our knowledge, there is limited study of rivers from SAR images for this region (Pai et al., 2019). The use of SAR image enables all weather monitoring of this ecologically sensitive region.
- Although the study area is the Western Ghats (Southern India), the experimental results demonstrate the cross-region generalization of the proposed approach. It is shown that the model trained for identifying rivers of one region (Southern India) can be utilized to identify rivers of a different region (Eastern India), thus creating an opportunity to develop a global river identification tool.

The manuscript is organized as follows: Section 2 presents the proposed approach along with the description of the study area and the dataset used, Section 3 discusses the results obtained using the proposed method for river identification and width measurement.

## 2. Materials and methodology

### 2.1. Study area and dataset description

This study focuses on measuring the river width of rivers flowing in the Mangalore–Udupi region of Coastal Karnataka (India). These rivers originate or pass through the Western Ghats, a UNESCO world heritage site recognized as one of the world's eight 'hottest hotspots' of biological diversity.[1] This study of river width would provide crucial inputs for ecologists and aid in water management of the region suffering from floods during monsoon (Service, 2020; Paniyadi, 2018) and frequent water scarcity in summer (Naina, 2019). Besides, the river width information could aid in increasing fish productivity, especially in capture-based aquaculture practices (Dineshbabu et al., 2012).

This study utilizes ortho-rectified SAR images with VH polarization, which is obtained from open source Sentinel-hub EO browser[2] for the Mangalore–Udupi region of Southern India. The SAR images were acquired for approximately three and half years (April 2017 to December 2020). In total, 45 images (one image per month) with a resolution of 1967x3004 pixels were obtained. Each pixel in the SAR images was then hand-labeled manually into two categories — rivers and non-rivers.

Following this, the pair of images (SAR image and labeled image) was padded with zeros at the right and bottom ends to get an image of dimension 2048x3072. Uniform crops of 256x256 were then taken from these images, which resulted in 96 sub-images from a single image. Fig. 2 shows the original SAR image along with two sub-images. Also shown are the corresponding ground truth labels for the sub-images. The river segmentation and river width measurement method presented next are applied to these sub-images. Note that the Mangalore–Udupi

[1] https://whc.unesco.org/en/list/1342/

[2] https://apps.sentinel-hub.com/eo-browser/

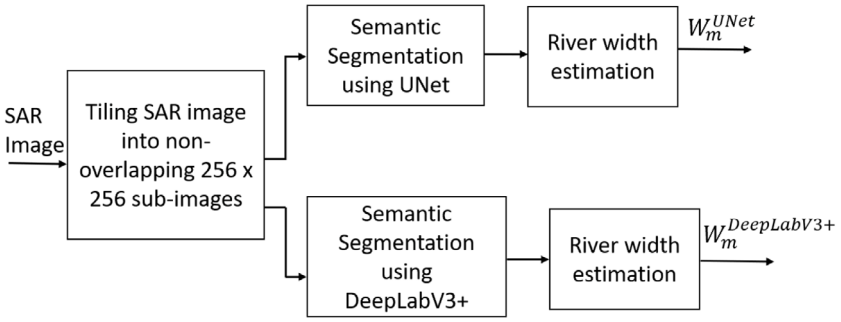
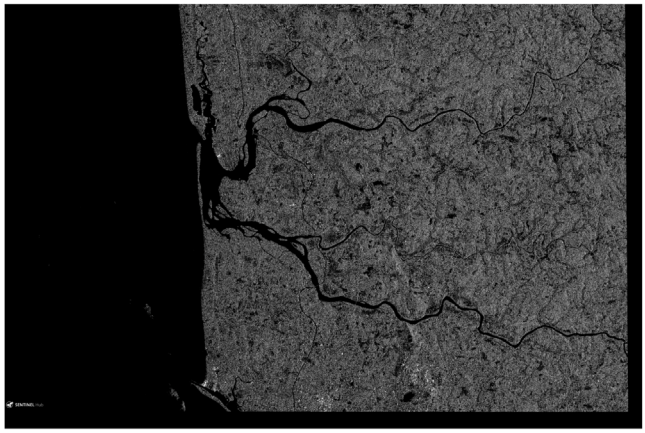
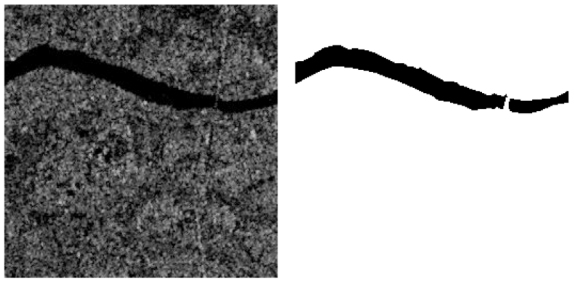


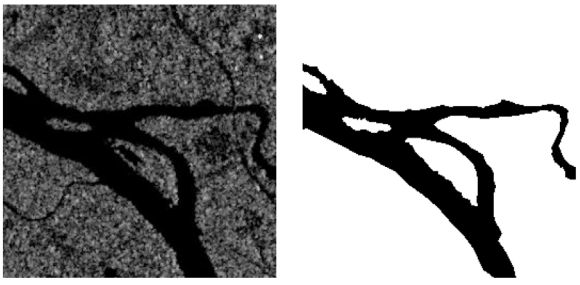

**Fig. 1.** Overview of the proposed approach for river width estimation: The original full size SAR image is split into 256 x 256 sub-images. The rivers are first identified in these sub-images using a semantic segmentation approach (U-Net or DeepLabV3+). Finally, the width of the river is measured from the segmentation output.



(a) Original SAR image.



(b) SAR Sub-Image (c) Ground truth mask



(d) SAR Sub-Image (e) Ground Truth Mask

**Fig. 2.** Dataset Preparation: Original SAR image of Mangalore–Udupi region along with two sub-images extracted from the original image. Also shown are the ground truth mask.

region receives heavy rainfall due to Indian monsoon from May–August every year. As a result, a few fields submerged with water can be seen in the SAR images of May and June 2019 (Fig. 6). This work focuses on estimating the river width, so these fields were not included as part of the ground truth mask.

In this approach, the rivers are first identified (Section 2.2), and subsequently, the width is measured from the segmented image (Section 2.3) with minimal user intervention.

### 2.2. River identification

The river identification is formulated as a two-class semantic segmentation problem where the foreground class contains all the pixels corresponding to rivers, and the background class contains the remaining pixels. Recently, Convolutional Neural Networks (CNN) based approaches have furnished state-of-the-art results on public benchmarks for semantic segmentation. A few of these algorithms include U-Net (Ronneberger et al., 2015), FCN (Long et al., 2015), DeepLabV3+ (Chen et al., 2018) etc. This work presents the performance evaluation of two state of the art semantic segmentation algorithms (U-Net, and DeepLabV3+) for river segmentation.

The availability of manually annotated data is an important factor in a supervised CNN based semantic segmentation methods. However, manually annotating each individual pixel in an image is a time consuming and error-prone task. The U-Net architecture is best suited for segmentation tasks with fewer training images (Ronneberger et al., 2015). Further, DeepLabV3+ is the state of the art method for semantic segmentation and has outperformed other methods on challenging semantic segmentation datasets such as PASCAL VOC, Cityscapes, and similar benchmarking segmentation datasets. Hence, these two methods have been considered in this study. A summary of these two approaches is presented below. More details about these methods can be found in Ronneberger et al. (2015), Chen et al. (2018).

A typical CNN based method for semantic segmentation consists of an encoder for extracting features from the images and decoder for inferring class labels for individual pixels. The encoder path of U-Net (Ronneberger et al., 2015) comprises four blocks of $3 \times 3$ convolution and max pooling operations. The first layer of this encoder consists of 64 feature maps with the number of feature maps increases by two after each layer. The symmetrical upsampling path again consists of four blocks. Each of these blocks contains a deconvolution layer, followed by a concatenation of feature maps from the corresponding encoder block. The result is then fed to $3 \times 3$ convolution layer with batch normalization. The encoder path extracts coarse contextual information while the decoder path helps in precise localization, thus obtaining an accurate segmentation map.

Another popular encoder–decoder based architecture used for semantic segmentation is DeepLabV3+ (Chen et al., 2018). This architecture is an improvement over the previous iteration DeepLabV3, which introduced parallel atrous convolution with different rates, thereby acquiring multi-scale contextual information. The encoder module of DeepLabV3+ is identical to the encoder module of DeepLabV3. The decoder module consists of a series of convolution and upsampling operations to produce an accurate segmentation map. Note that the

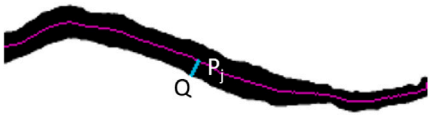



**Fig. 3.** The river centerline shown superimposed on the binary segmentation result obtained on a SAR sub-image. Also shown are the point on the river bank $Q$ and the river centerline $P_j$.

low-level features from the encoder path are concatenated before the convolution operation.

In this work, two semantic segmentation algorithms' performance is compared for river segmentation in SAR images. Note that it is possible that foreground pixels also contain other water-bodies (low lying fields submerged with water, lakes, wetlands, reservoirs, etc.) due to the similar reflection property of water in the SAR image. In such a scenario, a topological structural analysis is utilized on the binary segmentation map first to identify all the contours present in the image (Satoshi and Abe, 1985). Subsequently, the contour, which includes the maximum number of pixels (contour with the largest area), is selected as the river's region. Indeed, in a sub-image of resolution 256 × 256 pixels, it is highly likely that the contour with the highest area is a river.

### 2.3. River width measurement

The river's width is measured from the binary segmentation map consisting of the river pixels in the foreground class. The river width is measured as the distance between two points on the riverbank along the direction orthogonal to the localized centerline of the river. This is achieved in two steps: (1) computation of river centerline and (2) measuring the distance between two points on the riverbank along the orthogonal direction to the river centerline.

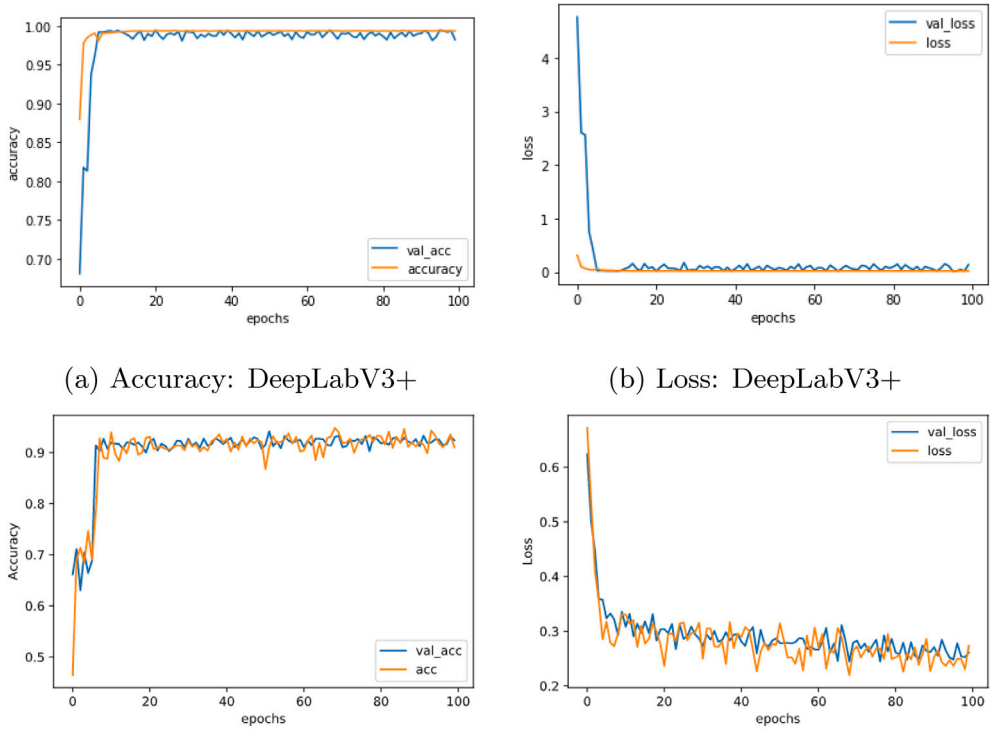
First, the medial axis (topological skeleton) of the foreground object is computed. The centerline of the river is then assumed to be represented by this skeleton. Let us represent $N$ pixels on the river centerline as $P_i, i = 1, 2, \ldots, N$. Given a point on river bank $Q$, the nearest point $P_j (P_j \in P_i)$ situated on the river centerline is first determined. The river width is subsequently computed as

$$W = 2 \;\; D_e(Q, P_j) \tag{1}$$

where $D_e(P, Q)$ represents the Euclidean distance between points $P$ and $Q$ (Fig. 3). This width $W$ is in the unit of the number of pixels, and can be scaled to the metric unit (Kilometers) by multiplying it with a scaling factor $\sigma$ : $W_m = \sigma W$. The scaling factor $\sigma$ is determined from the original SAR image.

### 2.4. Training procedure for river identification

The parameters for the two semantic segmentation algorithms (U-Net and DeepLabV3+) are learned on the SAR sub-image dataset. This dataset was split into training and test set based on the timestamp of the images. The training and validation split contains all sub-images from April 2017 till March 2019, which amounted to 2400 images. Note that the resolution of these sub-images is 256 × 256 pixels. The test split includes 2016 images, which were acquired from April 2019 to December 2020. Note that there are sub-images which contain only




(a) Accuracy: DeepLabV3+ (b) Loss: DeepLabV3+

(c) Accuracy: U-Net (d) Loss: U-Net

**Fig. 4.** Loss and Accuracy plot for U-Net and DeepLabV3+ models obtained on training (shown in orange) and validation set (shown in blue). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

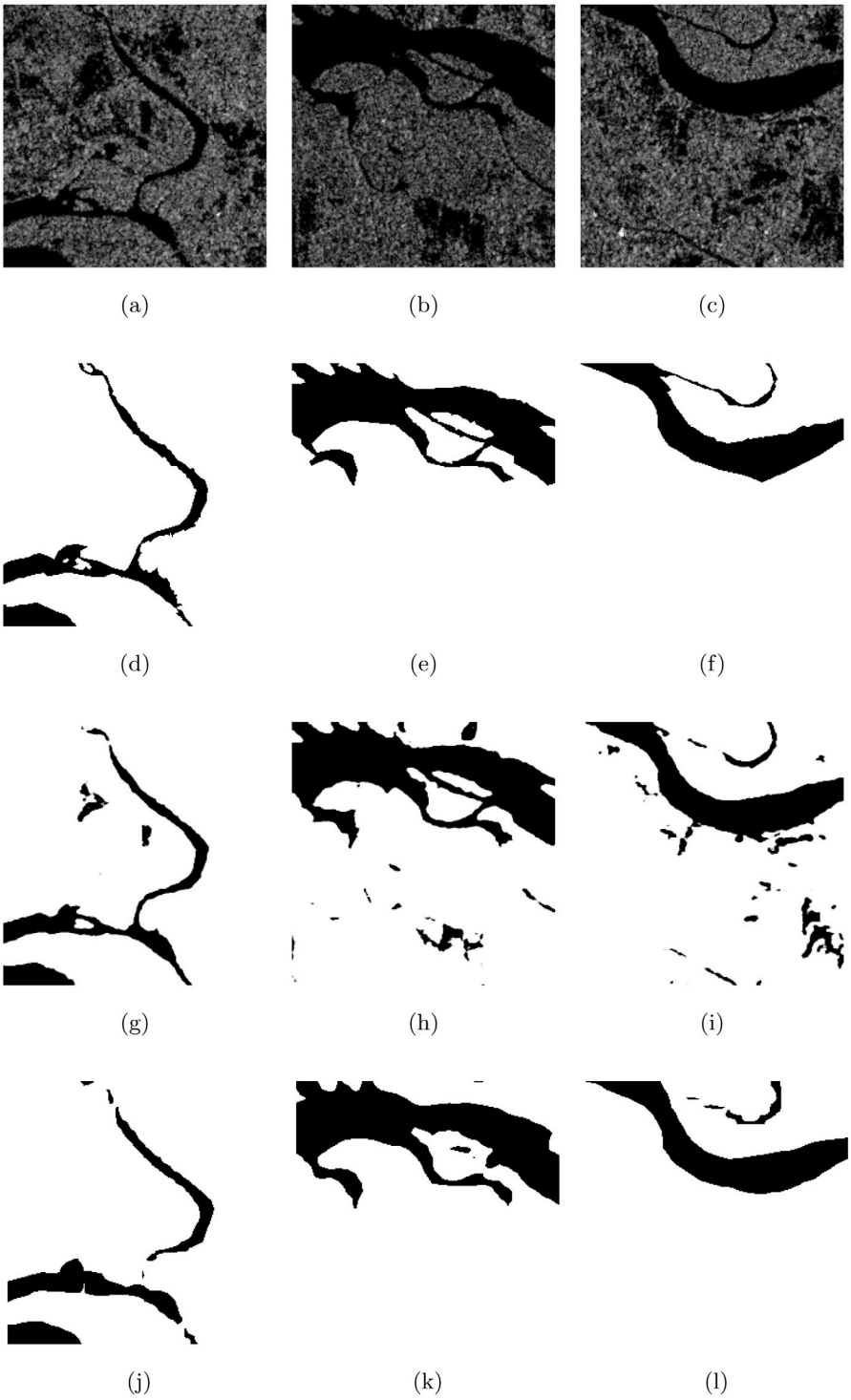



**Fig. 5.** Semantic Segmentation of SAR images: First and Second row shows the SAR sub-image and its corresponding ground truth. The third and fourth rows show U-Net Segmentation map and DeepLabV3+ segmentation map respectively.

land or only sea. Therefore, a random training–test split might create a situation with training/test split containing only land or only sea images. Thus, the dataset was split temporally into training and test split. The two semantic segmentation methods, U-Net and DeepLabV3+, are trained separately on the same dataset described above with the same training/test split. The FLOPs of the semantic segmentation module was : 6,20,49,671 for U-Net and 42,17,507 for DeepLabV3+. The models were trained and tested on a Nvidia GTX1060 mobile GPU with an Intel i7-7700HQ CPU @ 2.80 GHz × 8 processor. On an average, it took around 30–45 min to train the model.

### 2.5. Evaluation

*River identification.* The performance of semantic segmentation algorithms for river identification is evaluated by comparing it with ground truth images. The following metrics are computed by comparing the segmentation results and ground truth mask: mean Intersection over

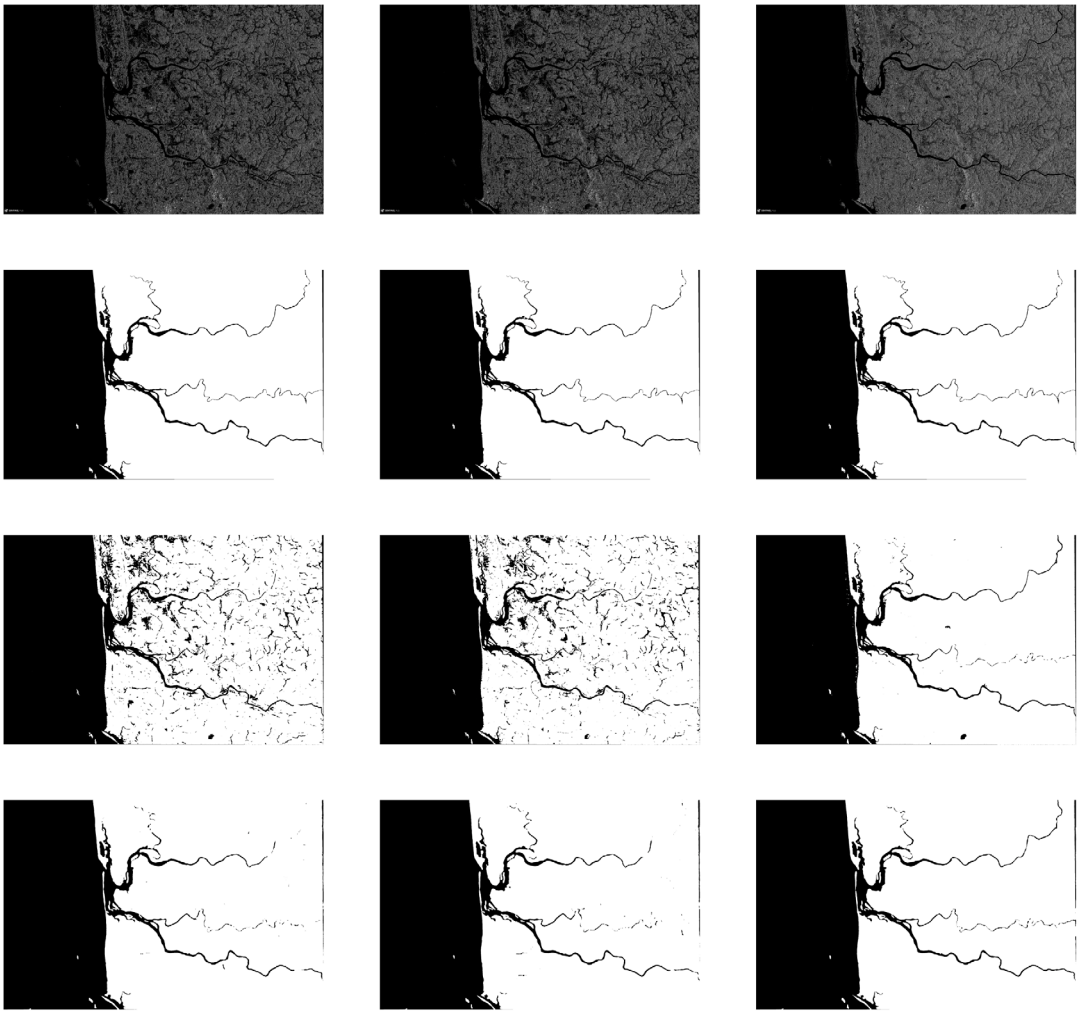


**Fig. 6.** Semantic Segmentation of SAR images. Top row represents the full size SAR image of May, June and September 2019 while the second row represents the corresponding ground truth. The segmentation results obtained using U-Net and DeepLabV3+ are shown in the third and fourth row respectively.

Union (mIOU), F1_score, and mean average precision (mAP). The mIOU and F1_score are defined as:

$$mIoU = \frac{\sum_i x_{ii}}{C\left(\sum_i \sum_j x_{ij} + \sum_j x_{ji} - x_{ii}\right)} \tag{2}$$

$$F1\_score = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{3}$$

where $C$ is the number of classes (two in this study), $x_{ij}$ represents the number of pixels belonging to class $i$ and predicted as class $j$, Precision $= \frac{TP}{TP+FP}$, Recall $= \frac{TP}{TP+FN}$ with TP, FP and FN are true positives, false positives and false negative respectively.

*River width measurement.* The river width estimated ($W_m$) using the proposed approach is compared with the actual river width ($W_a$) measured manually from Google Earth. The following metrics are computed to evaluate the performance of the proposed approach: Average Absolute Error (AAE) and Root Mean Square Error (RMSE):

$$Average\ Absolute\ Error = \frac{|W_m - W_a|}{N} \tag{4}$$

$$Root\ Mean\ Square\ Error = \sqrt{\frac{\sum_i (W_m - W_a)^2}{N}} \tag{5}$$

where $N$ is the total number of locations where river width is measured.

## 3. Results and discussion

### 3.1. River identification

This section compares the performance of the two semantic segmentation algorithms for river identification in SAR sub-images. The training was done for 100 epochs for both the U-Net and DeepLabv3+ with 50 steps per epoch. The loss function used was binary cross-entropy since this is a binary class segmentation problem. Adam's optimizer was used with a constant learning rate of $1 \times 10^{-4}$. The weights were initialized in accordance with Xavier Initialization (Glorot and Bengio, 2010). Fig. 4 shows the loss and accuracy plot for U-Net and DeepLabV3+ obtained on training and validation images of the SAR sub-image dataset.

The identification of rivers in these SAR images is challenging due to the presence of multiple water bodies and significant variation in the surface water (Fig. 2). The performance of U-Net and DeepLabV3+ algorithms are evaluated by computing mean Intersection Over Union (mIoU), mean Average Precision (mAP), and F1_score by comparing it with ground truth (Table 1). It can be seen that both the algorithms perform competitively and achieve an mIoU of 0.96. Fig. 5 shows the segmentation obtained using U-Net and DeepLabV3+. It can be observed that both methods can identify the river pixels. However, it was found that a more accurate river segmentation is obtained using U-Net as compared to DeepLabV3+ (Fig. 5). For instance, the stream after the river bifurcation is being identified by U-Net (Fig. 5(h))

**Table 1**
Evaluation Metrics for binary semantic segmentation of SAR images dataset described in Section 2.1.

| Metrics | U-Net | DeepLabv3+ |
|---|---|---|
| mIoU | 0.96 | 0.96 |
| mAP | 0.83 | 0.83 |
| F1 | 0.97 | 0.98 |

but DeepLabV3+ fails to accurately identify this stream (Fig. 5(k)). Besides, it can be seen that U-Net can identify the riverbanks more accurately as compared to DeepLabV3+. Further, the island situated in the river can be located using U-Net (Fig. 5(g)), but DeepLabV3+ erroneously segments it as a river (Fig. 5(j)). It is also observed that in a few images (less than 2% of test images), both the models (U-Net and DeepLabV3+) fails to identify very narrow rivers. For better visualization, the segmentation results on the full-size SAR image is shown in Fig. 6 for the entire region. The Mangalore–Udupi region receives heavy rainfall from May–August every year, primarily due to Indian monsoon. The three SAR images are shown in Fig. 6 were acquired in May, June, and September 2019. The low lying fields are inundated with water due to heavy rainfall and can be seen in the SAR images (as dark pixels) of May and June 2019. However, these fields are not filled with water in the SAR image of September 2019. These fields were not included as part of the ground truth mask, because this work focuses on estimating the river width. Surprisingly, U-Net was able to generalize better and could detect these submerged fields that were not in the training images.

These additional fields are ignored by analyzing the total number of pixels present in each waterbody using topological structural analysis. The region (waterbody) with the highest number of pixels in the SAR sub-image was selected, and this was indeed a river. For instance, Fig. 7 shows the segmentation results obtained using U-Net. It can be observed that the region with the highest number of pixels is indeed the river.

There are limited works on river identification by binary semantic segmentation (Pai et al., 2019, 2020) of SAR images to the best of our knowledge. In Pai et al. (2019, 2020), the performance of U-Net, FCN was studied for semantic segmentation of river in SAR images, and an mIOU of 0.95 (U-Net) and 0.91 (FCN) was obtained. However, the SAR image was resized to 512 × 512, to meet the input image dimension requirement of U-Net, and then segmented. This resizing would result in loss of details, thus affecting the accuracy of river width measurement. Instead, in the proposed approach, the semantic segmentation is performed on the sub-image, and the resulting segmentation map is restitched to obtain the segmentation of the full resolution image. The use of the SAR sub-image for segmentation also helps in identifying the river using topological structural analysis if multiple water bodies are detected.

The speckle noise is an undesirable effect of SAR image and there are several approaches for reducing the effect of speckle noise. In our study, it was found that the use of Lee filter for reducing the effect of speckle noise results in loss of image information. An mIoU of 0.90 was obtained with U-Net trained after removal of speckle noise with Lee filter, as compared to mIoU of 0.97 using UNet trained on images without filtering. Indeed, one of the challenges in despeckling SAR images is to suppress speckle noise while preserving image information such as edges. Therefore, the proposed approach does not contain a pre-processing step for removal of speckle noise.

### 3.2. River width measurement

The river's width was measured at 116 different locations in the SAR images from the period of April 2019–December 2020. The estimated river width was compared with the actual river width measured manually in Google Earth. Table 2 shows the measured width at few locations

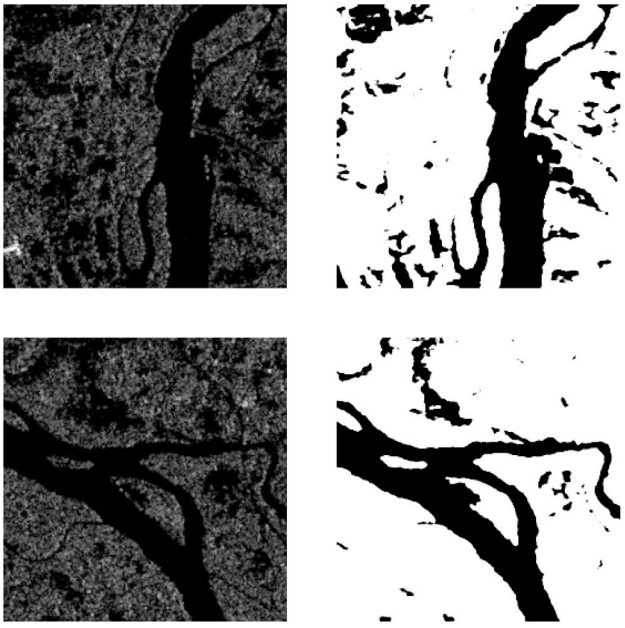


**Fig. 7.** Post-processing segmentation results for river identification: The segmentation obtained using U-Net might also contains low lying fields submerged with water (Left column). This results is further refined by assuming that the region with largest number of pixel represents the river. It can be observed that the rivers are indeed identified using this assumption (right column).

**Table 2**
River width measured using U-Net Segmentation results $W_m^{UNet}$ and DeepLabV3+ segmentation results $W_m^{DeepLabV3+}$.

| Actual width (Kms) | $W_m^{UNet}$ (Kms) | $W_m^{DeepLabV3+}$ (Kms) |
|---|---|---|
| 0.104 | 0.103 | 0.0749 |
| 0.117 | 0.1266 | 0.1286 |
| 0.17 | 0.1185 | 0.1499 |
| 0.192 | 0.159 | 0.118 |
| 0.059 | 0.0353 | 0.053 |
| 0.076 | 0.12366 | 0.05586 |
| 0.08 | 0.099 | 0.129 |
| 0.268 | 0.2529 | 0.2185 |

using the segmentation results obtained from U-Net and DeepLabV3+. The average absolute error in measurement was 20.05 m using the segmentation results of U-Net and 31.3 m using the segmentation results of DeepLabV3+. Besides, the root mean square error (RMSE) of 24.9 m using U-Net segmentation map and 76.6 m using DeepLabV3+ segmentation map was obtained. The estimated river width using the U-Net segmentation results is more accurate than that of DeepLabV3+, mainly due to a more precise segmentation map obtained using U-Net (Figs. 5, 6). This finding reinstates the importance of river identification to measure its width accurately.

Note that the proposed approach for river width measurement significantly outperforms the previously reported work. An average absolute error and RMSE were reported to be 43.1 m and 99.2 m respectively in Yang et al. (2020). In contrast, mean absolute error and RMSE of 20.05 m and 24.9 m is obtained in the proposed approach from the U-Net segmentation results.

## 4. Computer code availability

The source code used in this study is available at https://github.com/ArjunChauhan0910/DeepRivWidth.

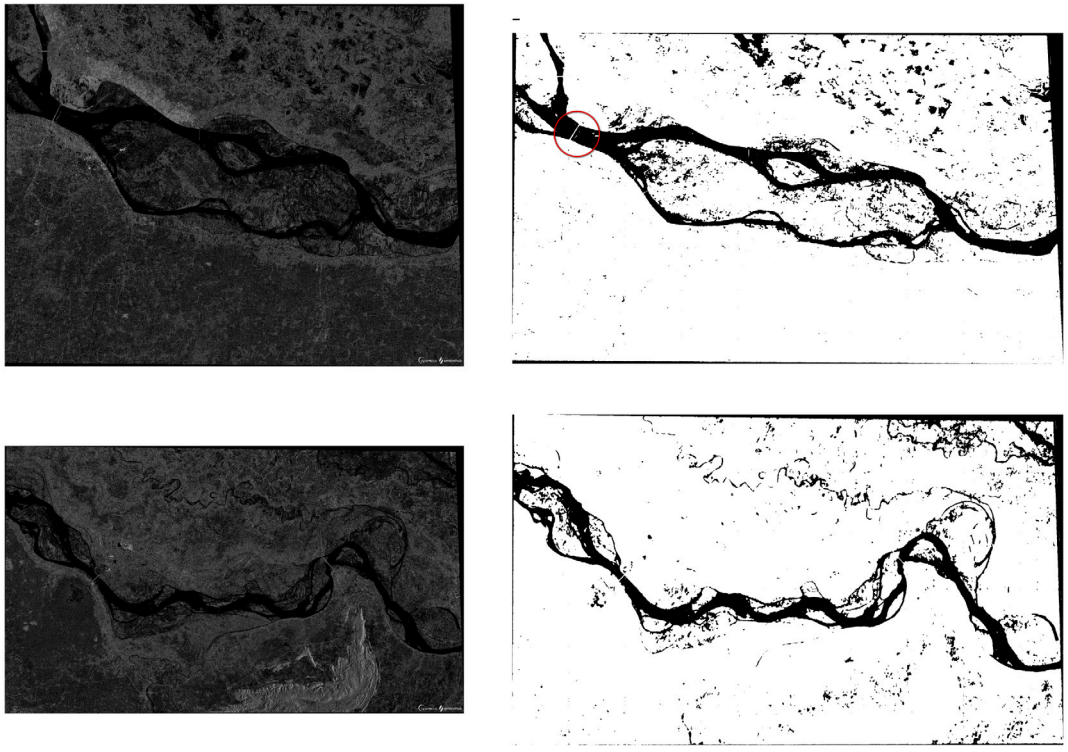


**Fig. A.8.** Original SAR image (left) of Ganges, Near Patna (Eastern India) and the semantic segmentation output (UNet) using the UNet model trained on images from Mangalore–Udupi Region (Southern India). It can be observed that the model trained on SAR images of one region can accurately identify the river of a completely different region due to the specular property of SAR images.

## 5. Conclusion

Typically, the river width is measured manually in the field which is a time-consuming and challenging task especially in rugged terrain. Remote sensing based methods offer an alternative approach to measure width remotely. The current methods use electro-optical sensors that are sensitive to obstructions, such as clouds. This work proposes to use Synthetic Aperture Radar images, which can penetrate clouds, for river width measurement. In addition, the SAR images allows an all-weather monitoring of the region and can even acquire the images even during the night. The existing methods of river identification in SAR images requires manual estimation of morphological filter parameters (Klemenjak et al., 2012), seed region (Ciecholewski, 2017) and thresholding parameters (Sghaier et al., 2016). In the proposed approach, the river identification is formulated as a binary semantic segmentation task and performance of two state-of-the-art methods for semantic segmentation is compared for river identification and subsequent width measurement. The use of end-to-end CNN network for river identification eliminates the need for manual estimation of morphological filter parameters, seed region and thresholding parameters. It is observed that the U-Net model could accurately identify rivers along with other water bodies. In fact, the U-Net model is able to segment the irregular shaped river banks (visible during Indian Monsoon) and small islands. Moreover, the proposed approach exploits the specular reflective property of the water in SAR image and formulates the river identification as a binary semantic segmentation task. Although the study area in this work is the rivers of Udupi–Mangalore region (Southern India), the experimental results demonstrate that the model trained on identifying rivers of Southern India can be used to identify rivers of Eastern India. This cross-region generalization paves the way for creation of a global river identification tool from SAR images. One of the limitations of the proposed approach is that in some instances, U-Net and DeepLabV3+ failed to identify a very narrow river. The absolute error between the estimated width and actual width is 20.05 m, demonstrating the proposed approach's robustness. The proposed method is validated by measuring the width of the rivers of the Mangalore–Udupi region of Coastal Karnataka (India), which is affected by frequent floods during monsoon and water scarcity in summer. The measured river width from SAR images would provide government authorities information for better planning and management of water resources of this ecologically sensitive region.

Indeed, the accuracy of river width measurement depends heavily on the robustness of the river identification method. This work compared the performance of two semantic segmentation algorithm for river width measurement. This comparative study helped gain more insights into these algorithms for the segmentation of rivers in SAR images. Despite the encouraging segmentation results, there is a scope for developing a more accurate river segmentation algorithm for SAR images. The dataset used in this study would be made available for further research in developing novel algorithms for river segmentation of SAR images and new application development.

## CRediT authorship contribution statement

**Ujjwal Verma:** Developed the method, Implementation, Result analysis, Contributed to the manuscript. **Arjun Chauhan:** Acquired data, Developed implementation, Result analysis, Contributed to the manuscript. **Manohara Pai M.M.:** Conceptualized the study and problem statement, Developed method, Result analysis, Contributed to the manuscript. **Radhika Pai:** Developed method, Result analysis, Writing - original draft.

## Appendix. Cross-region generalization of river identification approach

The motivation behind this work was to measure the river width of the ecologically sensitive Udupi–Mangalore region. However, the trained semantic segmentation model was also utilized to identify the

Ganges river, near Patna (Eastern India). Fig. A.8 shows the original SAR image of Ganges and the semantic segmentation output of the same region using the UNet model trained on SAR images of Udupi–Mangalore region (Southern India). It can be observed that the model trained on a particular region can accurately identify the river of a completely different region. Quantitatively, an mIoU of 0.84 and F1-score of 0.91 was measured for the two images shown in Fig. A.8. In fact, the trained model is able to precisely identify the minute details such as bridges (shown as red circle in Fig. A.8) across the river. This cross-region generalization is primarily due to the specular reflection property of water, and paves the way for a global river identification tool from SAR images.

## References


Chen, L.C., Zhu, Y., Papandreou, G., Schroff, F., Adam, H., 2018. Encoder-decoder with atrous separable convolution for semantic image segmentation. In: Proceedings of the European Conference on Computer Vision. ECCV. pp. 801–818.

Ciecholewski, M., 2017. River channel segmentation in polarimetric SAR images: Watershed transform combined with average contrast maximisation. Expert Syst. Appl. 82, 196–215.

Cordts, M., Omran, M., Ramos, S., Rehfeld, T., Enzweiler, M., Benenson, R., Franke, U., Roth, S., Schiele, B., 2016. The cityscapes dataset for semantic urban scene understanding. In: Proc. of the IEEE Conference on Computer Vision and Pattern Recognition. CVPR. pp. 3213–3223.

Dineshbabu, A., Thomas, S., Swathi Lekshmi, P., Sasikumar, G., 2012. Adoption of sustainable capture based aquaculture practices by traditional fishermen of karnataka. Indian J. Fisheries 59 (1), 49–52.

Girisha, S., Pai, M.M.M., Verma, U., Pai, R.M., 2019. Performance analysis of semantic segmentation algorithms for finely annotated new UAV aerial video dataset (ManipalUAVid). IEEE Access 7, 136239–136253. http://dx.doi.org/10.1109/access.2019.2941026.

Gleason, C.J., Smith, L.C., 2014. Toward global mapping of river discharge using satellite images and at-many-stations hydraulic geometry. Proc. Natl. Acad. Sci. 111 (13), 4788–4791. http://dx.doi.org/10.1073/pnas.1317606111.

Glorot, X., Bengio, Y., 2010. Understanding the difficulty of training deep feedforward neural networks. In: Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics. pp. 249–256.

Klemenjak, S., Waske, B., Valero, S., Chanussot, J., 2012. Automatic detection of rivers in high-resolution SAR data. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 5 (5), 1364–1372.

Li, W., He, C., Fang, J., Zheng, J., Fu, H., Yu, L., 2019. Semantic segmentation-based building footprint extraction using very high-resolution satellite images and multi-source GIS data. Remote Sens. 11 (4), 403. http://dx.doi.org/10.3390/rs11040403.

Li, X., Shen, H., Zhang, L., Zhang, H., Yuan, Q., Yang, G., 2014. Recovering quantitative remote sensing products contaminated by thick clouds and shadows using multitemporal dictionary learning. IEEE Trans. Geosci. Remote Sens. 52 (11), 7086–7098. http://dx.doi.org/10.1109/TGRS.2014.2307354.

Ling, F., Boyd, D., Ge, Y., Foody, G.M., Li, X., Wang, L., Zhang, Y., Shi, L., Shang, C., Li, X., Du, Y., 2019. Measuring river wetted width from remotely sensed imagery at the subpixel scale with a deep convolutional neural network. Water Resour. Res. 55 (7), 5631–5649. http://dx.doi.org/10.1029/2018WR024136, arXiv:https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2018WR024136, URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018WR024136.

Long, J., Shelhamer, E., Darrell, T., 2015. Fully convolutional networks for semantic segmentation. In: 2015 IEEE Conference on Computer Vision and Pattern Recognition. CVPR, IEEE, pp. 1–10. http://dx.doi.org/10.1109/cvpr.2015.7298965.

Mateo-García, G., Gómez-Chova, L., Camps-Valls, G., 2017. Convolutional neural networks for multispectral image cloud masking. In: 2017 IEEE International Geoscience and Remote Sensing Symposium. IGARSS, pp. 2255–2258. http://dx.doi.org/10.1109/IGARSS.2017.8127438.

Naina, J.A., 2019. Water crisis in Dakshina Kannada, Udupi. https://www.deccanherald.com/state/mangaluru/water-crisis-in-dakshina-kannada-udupi-728075.html. (Accessed 01 September 2020).

Pai, M.M., Mehrotra, V., Aiyar, S., Verma, U., Pai, R.M., 2019. Automatic segmentation of river and land in SAR images: A deep learning approach. In: 2019 IEEE Second International Conference on Artificial Intelligence and Knowledge Engineering. AIKE, IEEE, pp. 15–20. http://dx.doi.org/10.1109/aike.2019.00011.

Pai, M.M., Mehrotra, V., Verma, U., Pai, R.M., 2020. Improved semantic segmentation of water bodies and land in sar images using generative adversarial networks. Int. J. Semant. Comput. 14 (01), 55–69.

Paniyadi, G., 2018. Heavy rain, Udupi, Dakshina Kannada districts flooded. https://www.deccanchronicle.com/nation/current-affairs/080718/heavy-rain-udupi-dakshina-kannada-districts-flooded.html. (Accessed 01 September 2020).

Pavelsky, T.M., Smith, L.C., 2008. RivWidth: A software tool for the calculation of river widths from remotely sensed imagery. IEEE Geosci. Remote Sens. Lett. 5 (1), 70–73. http://dx.doi.org/10.1109/lgrs.2007.908305.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Lecture Notes in Computer Science, Springer International Publishing, pp. 234–241. http://dx.doi.org/10.1007/978-3-319-24574-4_28.

Satoshi, S., Abe, K., 1985. Topological structural analysis of digitized binary images by border following. Comput. Vis. Graph. Image Process. 30 (1), 32–46.

Service, E.N., 2020. Karnataka braces for floods as more rains expected. https://www.newindianexpress.com/states/karnataka/2020/aug/06/karnataka-braces-for-floods-as-more-rains-expected-2179637.html. (Accessed 01 September 2020).

Sghaier, M.O., Foucher, S., Lepage, R., 2016. River extraction from high-resolution sar images combining a structural feature set and mathematical morphology. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 10 (3), 1025–1038.

Song, D., Zhen, Z., Wang, B., Li, X., Gao, L., Wang, N., Xie, T., Zhang, T., 2020. A novel marine oil spillage identification scheme based on convolution neural network feature extraction from fully polarimetric SAR imagery. IEEE Access 8, 59801–59820.

Verma, U., Rossant, F., Bloch, I., 2015. Segmentation and size estimation of tomatoes from sequences of paired images. EURASIP J. Image Video Process. 2015 (1), 1–23. http://dx.doi.org/10.1186/s13640-015-0087-0.

Verma, U., Rossant, F., Bloch, I., Orensanz, J., Boisgontier, D., 2014. Shape-based segmentation of tomatoes for agriculture monitoring. In: Proceedings of the 3rd International Conference on Pattern Recognition Applications and Methods. ICPRAM 2014, pp. 402–411. http://dx.doi.org/10.5220/0004818804020411.

Wurm, M., Stark, T., Zhu, X.X., Weigand, M., Taubenböck, H., 2019. Semantic segmentation of slums in satellite images using transfer learning on fully convolutional neural networks. ISPRS J. Photogramm. Remote Sens. 150, 59–69. http://dx.doi.org/10.1016/j.isprsjprs.2019.02.006.

Yang, X., Pavelsky, T.M., Allen, G.H., Donchyts, G., 2020. RivWidthCloud: An automated google earth engine algorithm for river width extraction from remotely sensed imagery. IEEE Geosci. Remote Sens. Lett. 17 (2), 217–221. http://dx.doi.org/10.1109/lgrs.2019.2920225.

Zhou, L., Zhang, C., Wu, M., 2018. D-LinkNet: LinkNet with pretrained encoder and dilated convolution for high resolution satellite imagery road extraction. In: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. CVPRW, IEEE, pp. 192–1924. http://dx.doi.org/10.1109/cvprw.2018.00034.