A Survey: Extraction of River Networks from Satellite Images

Sukrut Bidwai, Devang Jagdale, Tejas Hiremath, Neil Bhutada, Sukhada Bhingarkar
Dr. Vishwanath Karad MIT – World Peace University, Pune, India
{s.bidwai2000, devangj9689, tejas.hiremath1, neilbhutada216}@gmail.com, sukhada.bhingarkar@mitwpu.edu.in

Abstract. River network extraction is crucial to keep track of the water resources. Various methods have been implemented in times series to yield profound and incisive outputs and are still being developed and combined with predefined available methods. We have carried out a structured survey on these methods and have presented them with their outputs. There are numerous websites available for data set collection. Some generalized methods are available like image processing, using pre-defined models, or developing user-defined algorithms. For image processing, various segmentation methods are available out of which clustering and threshold-based segmentation are mostly used. Pre-defined models such as CNN, ResUNet, YOLO, FasterCNN, MSCFF, etc. are available. These algorithms can be used for the extraction of River networks but might not yield higher accuracies. Hence, this paper concentrates mainly on approaches for the extraction of river networks from satellite images.

Keywords: CNN, Deep Learning, DEM, D8 Algorithm, Image-processing, River network Extraction, Mask Generation, Satellite Images, Segmentation.

1 Introduction

Water resource is the most significant source of life on Earth. Water resources monitoring, protection and flood management is a predominant factor. Extraction of Water Bodies from a satellite image is a crucial task. Previously the man-made space probes were assigned to collect huge volume of data in the form of graphic pictorial representations of certain surfaces, which resulted in load archives of images which were acquired. However, now this gigantic amount of data is reshaped into valuable executable data. Estimation and analysis of surface changes are crucial for a better understanding and management of processes that cause them. Various machine learning models such as Resnet, RCNN, YOLO, Faster-CNN, ResUnet, etc are further used for extraction of river networks. Initially the basic steps include pre-processing of images and converting them into uniform size for further training and feature extraction. Image spectrum and image enhancements are added in order to extract RGB & grayscale values which will be further trained to extract the exact river network from the provided image based on training and validation testing. Thus, from above research some attributes and models are taken into consideration for incisive outputs.

2 Data Source and Study Area

The data sources which are available consist of USGS, NASA, Google Earth Engine [1], Kaggle, Copernicus, Bhuvan NSRC GOV etc. There are various satellites which are available on the following websites e.g.: Sentinel-I [2], Sentinel- II [2], TERRA, LISS-III, etc. Images can be selected based on various features like using map tiles, interactive drawing, latitudinal and longitudinal coordinates etc. Some websites also provide random images from different parts of Earth. These regions include Heihe River Basin, KG river basin [3] which emanates southern area of tablelands. There are some websites available which provide filters for satellite images like shadows and clouds.

2.1 Image Processing techniques for Extraction of River Networks

Researchers proposed a method for river network extraction using image processing.

Process includes taking input as a Landsat satellite image and produces a river mask as output. Initially after taking satellite image as input, Enhancement using decorrelation stretching is performed. In this process image colours are enhanced by stretching their RGB bands which causes pixels to lose their correlation with each other which in turn causes them to normalize and removes high correlation to produce images with higher colour contrast. By doing this feature discrimination process becomes much easier [1].

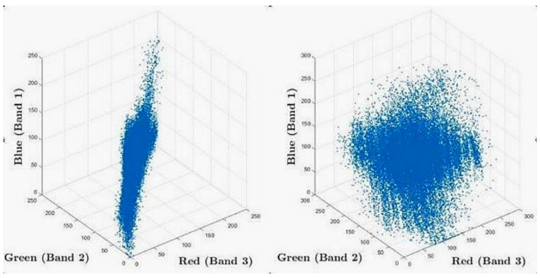

Fig. 1. Colour scatterplot before (a) and after (b) decorrelation [1]

After performing enhancement, the next step is image segmentation. Image is first converted to a grayscale image where grayscale level ranges up to 256 levels. Image segmentation is the process where we split pixels into various groups. Different groups contribute to different objects which makes similar pixels fall under one object and exhibit a common meaning. Image segmentation makes the image simpler which helps us to analyze the image at a much lower complexity [1].

There are four different types of segmentations which are available are as follows:

a) Threshold segmentation
b) Edge segmentation
c) Region segmentation
d) Clustering segmentation

Threshold segmentations are compared with the intensities of each pixel with a given threshold and assign that pixel a group. There are two types of thresholds available:

a) Global Threshold: There is only one common threshold for whole image
b) Local Threshold: There are multiple thresholds for different parts of images

Global threshold method was better suited for rivers mask as we had to only group pixels into two groups. In Threshold based segmentation, threshold value has to be selected appropriately. Best possible value can be obtained by trial- and-error method but instead of using an iterative method we opt for Histogram thresholding. Histogram Thresholding considers that a given image has two parts: subject and background. It tries to find an optimal threshold value(T) so that the image can be divided into subject and background. If (image[x][y] <= T) assign group G1 Else assign group G2 In above example x and y denote intensity of pixel at given coordinates. In this way we can assign pixel groups where G1 and G2 act as object and background. In Clustering based segmentation we try to group pixels that are together. Pixel which are closer to each other are considered to be in relation and are grouped together. We can use K-means clustering for segmentation. In clustering we try to find the cluster centre so that pixels close to that centre will be assigned a cluster. We keep repeating this process until we have only two clusters, subject and background [1].

Region based segmentation is a repetitive process where pixels are grouped into smaller regions in each iteration to form a larger region. Initially we define some seed points and from these points our region starts growing by claiming neighboring points which have similar properties like colour, intensity, grey scale level, etc. This method has high time complexity compared to other segmentation methods and it is also sensitive to noise.

Segmentation gives output as a binary image where the object is river(white) and background is land(black). Final step is to remove clutter from the image. After segmentation, the image might have noise or unwanted objects categorized as subject. To remove these unwanted objects from the image we use morphological operations. In morphological operations we compare a given pixel with its neighbor and decide whether to keep it or not. Process of removing pixels from an image is called Erosion. In this process we initialize the number of neighboring pixels a pixel should have. Pixel having fewer neighbors than selected value is transferred to another group i.e., transfers pixel from subject to background. In this way we obtained a river mask for a Landsat satellite image.

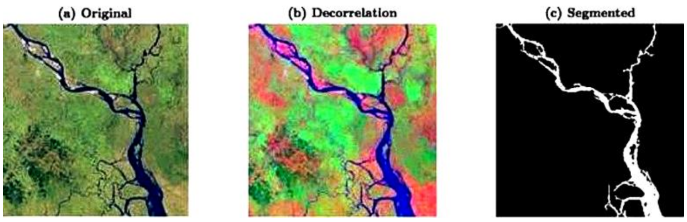

Fig. 2. Image segmentation for satellite images [1]

The only limitation is that this method is not generalized. For images with different features like snow, clouds or shadows might result in false mask prediction.

2.2 CNN Method for Extraction of River Networks

The river networks can be extracted using few methods. One of them is using a predefined model and second would be developing a new model based on any predefined model and using their outputs as new input to another model. Using a predefined model does not yield high accuracy at all times so instead of using predefined models we can combine different models to get high accuracy. One of the methods proposed includes a combination of multi spectral data and SAR (synthetic aperture RADAR) data to jointly map the extent of the river over large regions. A neural network structure is defined to extract common information from multi source data inputs i.e.: multi spectral data and SAR (synthetic aperture RADAR) data [2].

A CNN is used to extract feature 1 and feature 2 from a network which is blocked and would be used during the model training and parameter transformations. In [2] Algorithms which have been defined include Algorithm 1 and Algorithm. Whereas Algorithm 1 shows the pseudo-code for the model training. Here, the training set with two different types of the sample pairs which are marked with a flag variable. Here, the flag variables are manually obtained during the multi-source sample pairs creation for common information extraction model training. the training sample pair with flag = 1 represents the relevant pair (ai, bi) and the training sample pair with f lag = 0 represents the irrelevant pair (ai, bj). Given two input samples Input 1 (e.g., ai) and Input2 (e.g., bi or bj), the loss function for extracting common information has been defined as:

$$\text{Loss (Input1, Input2, f lag)} = \text{Loss flag=0} + \text{Loss flag} = 1 \quad (2)$$

When flag = 0, the Loss flag=0 is computed as follows:

$$\text{Loss flag} = 0 = 0.5 * (1 - \text{flag})\ \{\max(0, m - \text{Distance})\}^2$$

In [2] Algorithm 1 defines the structure as the input acquired are a sample pair of SAR and multi spectral data, and a flag which is further defined as a multi-source training sample pair of relevant or irrelevant type. The output derived from this algorithm 1 is the distance variable. Now the feature 1 and feature 2 which are extracted from the CNN are blocked and the Distance which calculated above is enabled with the combination of input 1 and input 2 i.e.: SAR and multispectral image sample input. The next step defined resembles of Fine-tuning the network parameters using loss function Loss (Input1, Input2, flag) defined in Equation. For each multisource training sample pair.

A new parameter is defined as m now here M, is calculated as the Euclidean distance between input 1 and input 2 as given by Distance = || f(input1) || ^2.

In [2] Algorithm 2 defines the structure for testing the process of common information extraction function. Here the feature 1 and feature 2 which were extracted using a CNN are enabled and the distance parameter is blocked. Input 1 and input 2 which are SAR data and multi spectral data source respectively are considered as input. A common information matrix for multi-source testing sample pairs are produced from feature 1 and feature 2 for each multi source testing sample pair.

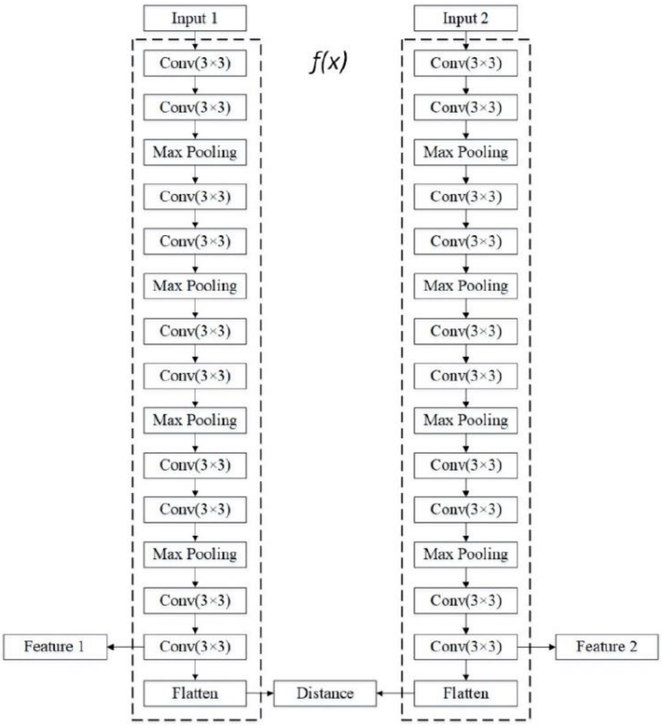

Fig. 3. CNN model for extraction of Feature 1, Feature 2 and Distance parameter. [2]

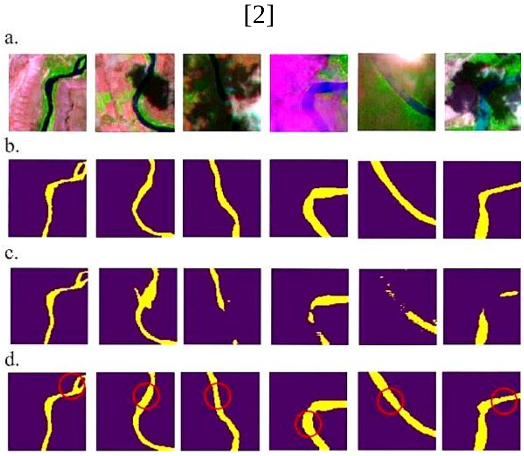

Fig. 4. Images in (a) are satellite images, images in (b) are ground truth, (c) output generated from pre-defined UNET model and images in (d) proposed method output. [2]

2.3 Mapping Network using Geo-Processing:

Digital elevation model consists of 3D characterization of elevation to represent terrain. It represents the fluctuating values between 2 points of elevation. The common use of DEM is to extract the terrain parameters for geomorphology. Three types of DEM (SRTM1-V3, AW3D, GDEM2, ALOS PRISM) are used to compare the changes in morphology. The Alos-prism module data is used to create the DEM. Demoted values are first identified and then filled in the DEM [3].

In [3] the D8 algorithm follows a basic principle that each cell or each grid can have eight different possible flows based on the maximum gradient. The D8 consists of mono flow parameter algorithm and leads with a principle which states that only 8 possible flows can be implemented in a Single grid.
The algorithm exhibits further formula:

$$\text{Single Grid} = H \div C$$

Where, H = The height difference between the two cells and C = The distance between centers of two cells.

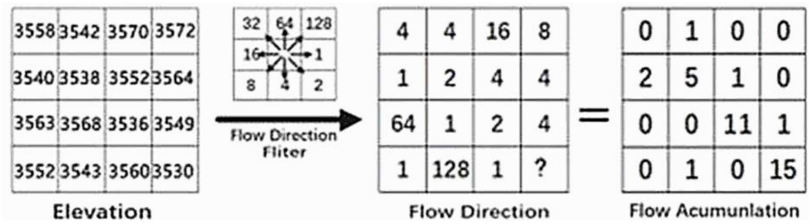

Fig. 5. Basic Conceptual model of D8 algorithm. [3]

The flow direction is determined using the principle of D8 algorithm. These directions are expressed by 8 different signatures 128, 1, 2, 4, 8, 16, 3. The D8(eight direction) flow direction method assumes that water from any cell will flow to one of the steepest neighboring cells. After calculating the direction of flow from D8, furthermore calculation of the upstream confluence matrix is carried out and then model extracts the critical support areas from the raster map of classified water body network. Hence, raster diagram segmentation the river network is extracted [3].

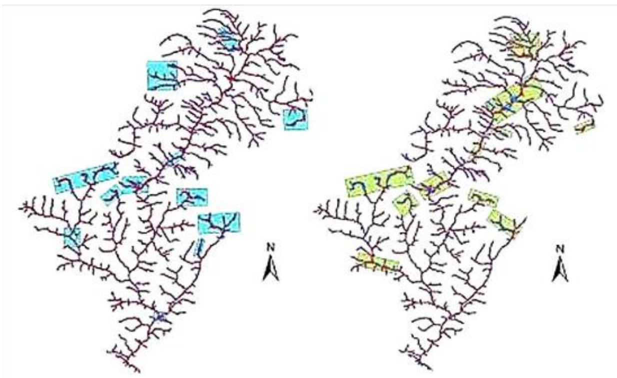

Fig. 6. The contrastive image of Drainage system in SRTM1v3, ALOS3D and ALOS prism algorithm [3]

Table 1. Summary of Existing Approaches

Sr No.	Methodology	Advantages	Limitations
1.	Enhancement of images using the correlation, image segmentation, clutter removal [1]	It generates very high accuracy mask compared to predefined models	The parameter set which was used within the algorithm were not generalized for image segmentation and noise data elimination
2.	Contrastive learning process to extract representative hidden features from multi- spectral data and SAR data [2]	Combined two different predefined models to produce river mask	2 datasets used, 1. Extremely Low-res images. 2. Noisy dataset.
3.	The basic structure is divided into bi-parts: a) Offline model training with MS CNN b) Online prediction on Google Earth Engine [5]	Provides an executable deep learning models with respect to Google Earth Engine to improve the accuracy of River extraction	Image dataset was only of 36 images
4.	Feature extraction using Gabor filter and canny edge detector are implemented for training and testing the random forest models [6]	Consistency is achieved which results in accuracy for network extraction	Only Canny's algorithm was used as an edge detection algorithm for feature extraction.
5.	A network is proposed, labeled as dense local feature compression (DLFC) network which aims to extract water body from different remote sensing images [7]	The experimental results are observed in which the model is perfectly able to distinguish between building and shadows (noise)	Distinctively small water bodies over certain set image dimensions are mixed in datasets.
6.	(CNN) is proposed, based on the multidimensional densely connected CNN for identifying water bodies in the lake area [8]	Optimal performance is achieved in identifying the boundaries and edges between land & water, whereas the mountain shadows and noise are distinguished	Optimal threshold to extract water varies with region and time
7.	ResUNet is used as backbone in the architecture, along with different post processing methods like watershed methods [9]	Automatic river extraction is been done using predefined CNN algorithms.	Dataset was collected in a small duration
8.	Identification and removal of noise using suitable filters using k-means & region growing segmentation techniques for river extraction [10]	MSE are found to be lowest for k-means clustering algorithm compared to region- based clustering.	Climatic conditions and shadows were not taken into consideration.

3. Conclusion

Image processing gives accurately predicted river masks for satellite images. But this method is applicable for only one type of image. If the images contain a wide spectrum of filters like clouds, shadows, snows, etc. it does not adapt and applies an iterative method for all. If we have all images of the same kind then image processing might be an appropriate choice. If we use predefined CNN models, we might get masks with lower accuracy as it is generalized and not specifically developed for river network extraction. But as future scope we can develop ensembled predefined CNN models or tweak some of its parameters to get higher accuracy. Also, an additional benefit of diverse image type can yield and support machine learning algorithms to yield a higher accuracy in this way we can use these methods for extraction of river networks to keep track of water resources through satellite imagery.

References

1. Ciaburri, Carey & Benitez, Monica & Sheta, Alaa & Braik, Malik. (2020). Automatic extraction of rivers from satellite images using image processing techniques. ACCENTS Transactions on Image Processing and Computer Vision. 6. 32-41. 10.19101/TIPCV.2020.618040.
2. Wei, Zhihao, Kebin Jia, Pengyu Liu, Xiaowei Jia, Yiqun Xie, and Zhe Jiang. 2021. "Large-Scale River Mapping Using Contrastive Learning and Multi-Source Satellite Imagery" Remote Sensing 13, no. 15: 2893. https://doi.org/10.3390/rs13152893
3. Chen, Feiyu & Tian, Bingwei & Adhikari, B. & Gou, Xiaoyun. (2019). Mapping Digital Drainage Network Using Geoprocessing: A Case Study of Kali Gandaki River Basin, Nepal Himalaya. 10.1109/IGARSS.2019.8900474.
4. Yang, Xiao & Pavelsky, Tamlin & Allen, George & Donchyts, Gennadiy. (2019). RivWidthCloud: An Automated Google Earth Engine Algorithm for River Width Extraction from Remotely Sensed Imagery. IEEE Geoscience and Remote Sensing Letters. PP. 1-5. 10.1109/LGRS.2019.2920225.
5. Asokan, A., & Anitha, J. (2019). Machine Learning based Image Processing Techniques for Satellite Image Analysis -A Survey. 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon). doi:10.1109/comitcon.2019.8862452
6. Wang, Yudie & Li, Zhiwei & Zeng, Chao & Xia, Gui-Song & Shen, Huanfeng. (2020). An Urban Water Extraction Method Combining Deep Learning and Google Earth Engine. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 13. 768-781. 10.1109/JSTARS.2020.2971783.
7. Zeba Naaz , Dr. G. Malini Devi, 2021, River Network Classification from Multi-Spatial Satellite Imagery using Random Forest, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) Volume 10, Issue 09 (September 2021),
8. Li, Mengya & Penghai, Wu & Wang, Biao & Honglyun, Park & Hui, Yang & Yanlan, Wu. (2021). A Deep Learning Method of Water Body Extraction From High Resolution Remote Sensing Images With Multi-sensors. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. PP. 1-1. 10.1109/JSTARS.2021.3060769.
9. Wang, Guojie, Mengjuan Wu, Xikun Wei, and Huihui Song. 2020. "Water Identification from High-Resolution Remote Sensing Images Based on Multidimensional Densely Connected Convolutional Neural Networks" Remote Sensing 12, no. 5: 795. https://doi.org/10.3390/rs12050795
10. Waldner, F., & Diakogiannis, F. I. (2020). Deep learning on edge: Extracting field boundaries from satellite images with a convolutional neural network. Remote Sensing of Environment, 245, 111741. doi:10.1016/j.rse.2020.111741
11. Deepika, R. G. M., & Kapinaiah, V. (2017). Extraction of river from satellite images. 2017 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT). doi:10.1109/rteict.2017.8256591