Extraction of River Networks from Satellite Images using Image Processing & Deep Learning Techniques

Devang Jagdale
Department of Computer Science and Engineering
Dr. Vishwanath Karad MIT – World Peace University, Pune
Maharashtra, India
devangj9689@gmail.com

Sukrut Bidwai
Department of Computer Science and Engineering
Dr. Vishwanath Karad MIT – World Peace University, Pune
Maharashtra, India
s.bidwai2000@gmail.com

Tejas Hiremath
Department of Computer Science and Engineering
Dr. Vishwanath Karad MIT – World Peace University, Pune
Maharashtra, India
tejas.hiremath1@gmail.com

Neil Bhutada
Department of Computer Science and Engineering
Dr. Vishwanath Karad MIT – World Peace University, Pune
Maharashtra, India
neilbhutada216@gmail.com

Sukhada Bhingarkar
Department of Computer Science and Engineering
Dr. Vishwanath Karad MIT – World Peace University, Pune
Maharashtra, India
sukhada.bhingarkar@mitwpu.edu.in

***Abstract* - River networks are widely observed and scrutinized for various purposes, which incorporate determining the terrestrial positions of water bodies, examining the gauge levels of the river, predicting river flows, and conserving sustainable energy resources as a consequence of Global warming. Extraction of these River networks on digital imagery systems are executed by various segmentation and machine learning model integration. In this paper, distinct datasets are used from Kaggle and Google Earth Engine, Segmentation methods such as Image segmentation, gray scaling, enhancement, global thresholding, and Deep Learning UNet Architecture are integrated with contemplation of extracting river networks from satellite images which result in achieving 80.98 % dice score for the developed UNet Model. Hence, these developed techniques can further be used for river extraction from satellite images. And can be applied to various semantic segmentation detection datasets.**

Keywords - Global thresholding, Image Processing, Mask, River network extraction, Satellite images, Segmentation, Unet Architecture.

I. Introduction

Water bodies are considered a prominent resource over the globe. To preserve mankind, these resources are at a predominant level. Most of the vegetation or agricultural profession requires water as its base element. Although, extracting a river network can be accomplished by traditional tasks too which include manually measuring the field width and by measuring river basins topographically from satellite images or graph maps. But both of these techniques tend to consume costly error-prone accuracies in it. Mapping a network from a digital image is a critical task to execute. Moreover, mapping a network from a satellite image tends to execute an over-designed supervised or unsupervised machine learning algorithm. Extracting a river network is a significant strategy to keep a track of water bodies over a timestamp and supervising such a track record can be beneficial for geologists and environmentalists to plan succeeding global changes in urban and rural areas over the globe. Diverse approaches have been designed and developed by developers to yield an accurate network map. Data sources are abundantly available but sorting out the required correct data requires manual sorting for now. Few methods are used for extracting such a network followed by, image processing using various segmentation methods such as clustering and threshold-based segmentation, using Machine learning models such as FCN, PSPNet, and SegNet accessible at a community level. In this paper, the image processing and U-Net model are engaged to pre-process, train, test, and validate the required image with its subsequent outputs. Hence, this paper purely emphasizes extracting and generating a binary mask from a river network and a water body from satellite images using image processing and deep learning technique.

II. Related Work

Semantic segmentation is a renowned method where research has been performed on the subject explicitly, segmentation on the satellite images is accomplished by very few researchers. In [2] paper, a contrastive method has been proposed where researchers have built a multi-source data segmentation model in which the model extracts common information from different data sources simultaneously preserving distinct knowledge from those datasets. F1-score, Recall, and Precision methods were used for metrics evaluation. The authors in [11] paper used the clustering method for mask generation where k-means and region growing clustering methods are used. They conducted their research on a certain part of the river, where the best result with the maximum Peak signal-to-noise ratio came out to be 72.33 with a Mean squared error of 0.0039 for the Niger River delta. In [9] paper a water identification from high-resolution satellite images was carried out where researchers proposed a CNN model based DenseNet and compared its result with other DeepLearning models like SegNet, ResNet, VGG, e Normalized Difference Water Index, and DeepLab v3+. The research was conducted on a lake named Poyang Lake. Metrics used for evaluation were Precision, Recall, F1 score, and mean Intersection over Union (MIOU). Finally, in the [12] paper UNet model was implemented along with ResNet for building segmentation. Our UNet model is partially inspired by this model. In the paper authors' model was dedicated to segmentation of building from satellite images where they used MIOU for metrics evaluation.

Authors achieved accuracy ranging from 0.79 to 0.84 with 100 epochs.

III. Dataset

Assortment of the dataset is a vital part of getting superior accuracy. Government and private entities which provide satellite image datasets are abundantly available over the internet. Some of them are ISRO Bhuvan NRSC, USGS.gov, Google Earth Engine Copernicus, etc. For the UNet model research dataset was collected from Kaggle and google earth engine where the Kaggle dataset was used for deep learning which comprised of a pair of satellite image and masks whereas google earth engine images was used for image processing. The size of data for image processing was only 30 images as size did not contribute to more accuracy. Dataset obtained from Kaggle comprised 2841 images out of which some images were discarded during the manual inspection as images lacked set guidelines. For the final prediction, after testing on the test dataset a required small dataset of 20 images were randomly included from google earth of various terrain consisting of a river network.

IV. Methodology

A. *Image Processing:*

In image processing, OpenCV was used to perform semantic segmentation using the global thresholding method. To get the best results enhancement technique are used where the gamma correction method to increase the brightness of images is incorporated. Gamma correction is a technique where conversion takes place for every pixel from 0 - 255 range to 0 - 1. Below formula is applied to each pixel.

$$O = I^{(1/G)}$$

Here I is input image and G is gamma value. By proceeding, O is converted back to its original range. If the gamma value is below 1 the the value is scaled down and if it is greater than 1 then the value is upscaled. If gamma value is exactly 1 then it remains unchanged. This method is implemented using opencv function which is a Look up table (cv2.LUT (inputImage,table)). After increasing brightness,RGB images are converted into grayscale images. The next step in process is to apply a thresholding algorithm where the final output i.e. mask will be generated. Global thresholding is one of the easiest algorithm where global threshold is defined for an image and if given pixel value is less than or equal to it then a 0 is assigned and if greater than, then 1 is assigned respectively. Consider (x,y) as coordinates for given pixel in a image and T as global threshold then,

$$if\big((x,y) > T\big)threshold(x,y) = 1$$

$$else\ threshold(x,y) = 0$$

For implementing this method OpenCV function was used which is cv2.threshold(). While defining the global threshold instead of keeping a common value for all images, the mean of that image was used as a global threshold. Mask for respective images were generated as a binary image where object (river) was represented as white and background (non river part) as black but, for the difficult images where color difference between the land and the river were less, mountain shadows, clouds, snow, etc were present in it. Hence, such images were falsely classified as a river network.

B. *UNet*

a) Architecture

Unet is one of the most widely used deep learning

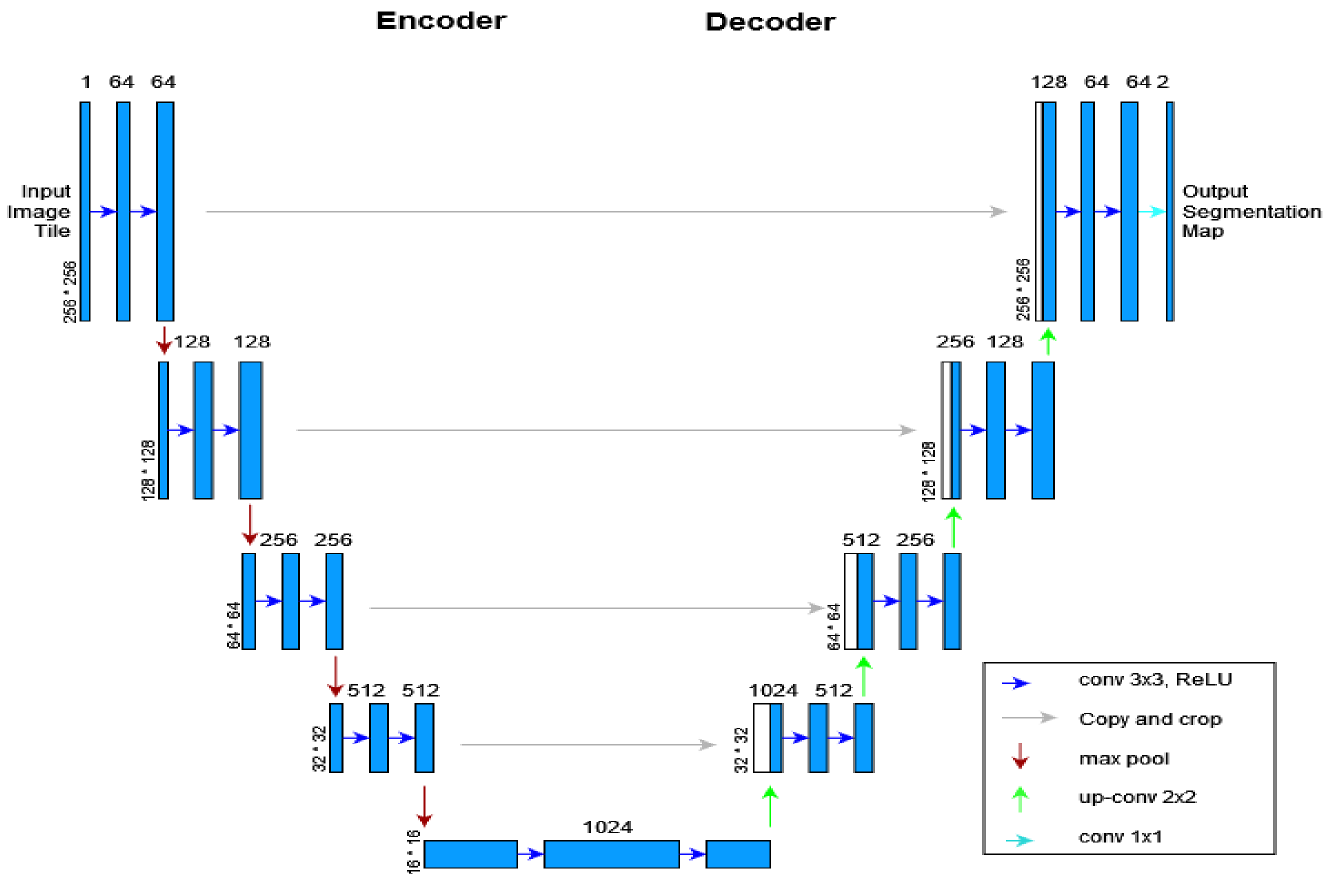

Fig: UNet Architecture

algorithms developed for semantic segmentation. It was developed for medical image segmentation and after obtaining great results later it was also incorporated into various image segmentation models. Unet comprises of two parts where first is the encoder which is used to extract features from input images where it uses down sampling to reduce image size and extract features. The second part consists of decoder which uses these extracted features as input to generate an output image. It uses upsampling and concatenation to generate output. It uses features which are extracted during the encoding process. In our model, slight modifications have been applied to the original model. Instead of using pre-trained weights, training of the model was carried out from scratch as the original model was not trained using satellite image data. For training, images were converted into (256 * 256) dimensions. Instead of using default cross-entropy for loss evaluation, diceBCEloss was used which is the sum of dice coefficient and BCE. Dropout was skipped over as the model did not encounter any signs of overfitting.

b) Metrics

For Semantic Segmentation Dice coefficient metrics is one of the most well-known metrics for model evaluation. The dice coefficient is calculated as.

$$Dice = \frac{2 * |A \cap B|}{|A| + |B|}$$

Where |A ∩ B| is the Area of Overlap and |A|+|B| is the total number of pixels in both images. Dice coefficient ranges between 0 to 1 where 1 denotes perfect overlap between images. It is very similar to IOU(Intersection-Over-Union) method.

c) Data Processing

Initial proceedings include a manual inspection, scarce amount of satellite images was blurred or partially out of frame due to which such images were eliminated from dataset. Similarly, some masks were falsely predicted which were also eliminated during manual inspection. Where satellite images were resized to a size of (2562563) and masks to (2562561) dimensions respectively.

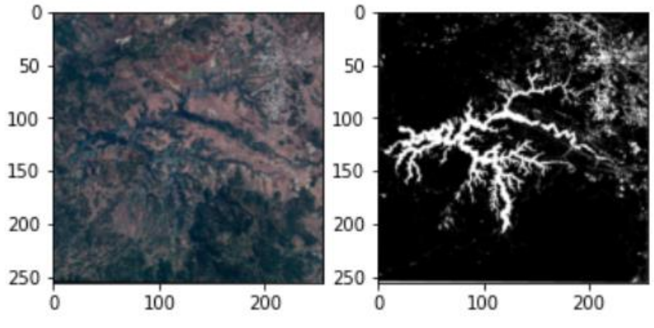

Fig: Satellite image with its corresponding Ground Truth Mask

Dataset was split as 80% for training and 20% for testing. Later data augmentation was applied as the dataset was not adequate for training and also shuffled the data. A pictorial representation of a pair of satellite images and masks are depicted below for reference.

d) Training Parameters and Environment

Google Colab was used for training and testing purposes as it resembles powerful computational tool with minimum charges. While training, number of the epochs were set to 12 and the batch size was 8. The time to train the model required 120 minutes (Approximation) where the learning rate was set to 1e-4. While training GPUs were used which are inbuilt GPUs named Tesla T4 provided by colab. Cuda can be accessed using an inbuilt torch function to save time on training. All the datasets were stored on google drive preferably as it resides within same environment with ease of access. Libraries used for model building and training were torch, NumPy, and OpenCV, and for visual presentation and plotting, matplotlib and tqdm were used. Gradio was used to create a frontend so that it can be incorporated into a website for further testing and deployment.

V. Results and Discussion

Output generated using image processing appears to be superior and edges were also sharper compared to the deep learning model. But the image processing technique parameter such as global threshold is altered for every image. Depending on image quality and brightness mean value of the image is not always the best threshold for the mask generation. For images with less color saturation between river and land part, the mask picks up some of terrestrial part as a river. Images with decent brightness and color separation of river and land image, processing technique do perform mask generation decently. This technique lacks to fulfil a generalized approach for satellite image. Model dice coefficient score is 80.98 which was trained for only 12 epochs. The Model predicts random images with a decent accuracy which are portrayed below. Predicted images are acceptable but random noise generation in images are observed. Edge of water bodies in images are slightly blunt, compared to the ground truth images in the areas where noise exists. But overall predicted mask of the satellite image is clearly indicating the river and small networks present in it. The below image represents the diceBCELoss graph for U-Net model.

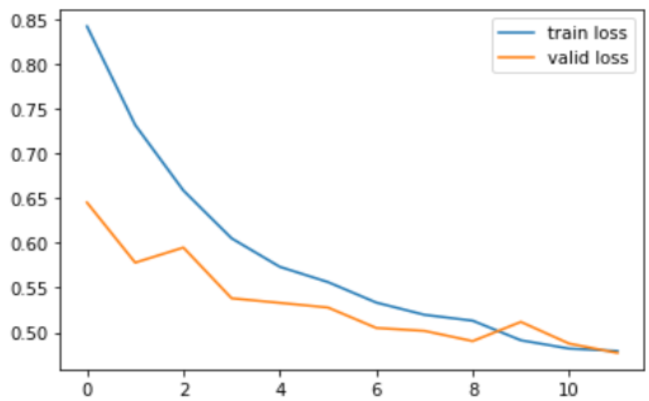

Fig: Training Vs validation Loss for UNet model

A. *Outputs and Screenshots*

a) Image Processing

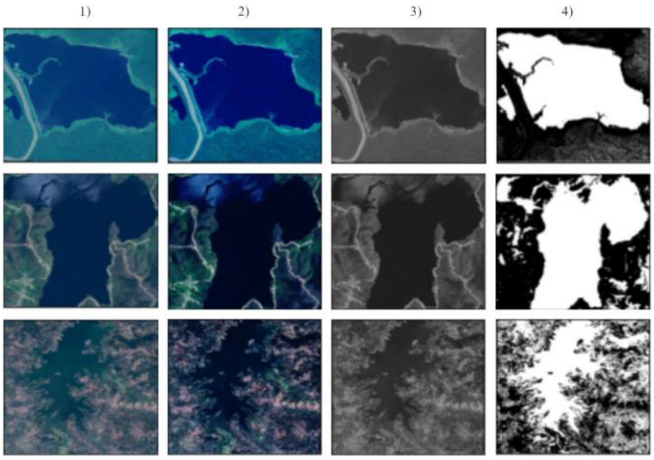

Fig: In above image 1) column refers to original image, 2) column refers to enhanced image, 3) column refers to grayscale image, and 4) column refers to Mask

b) UNet

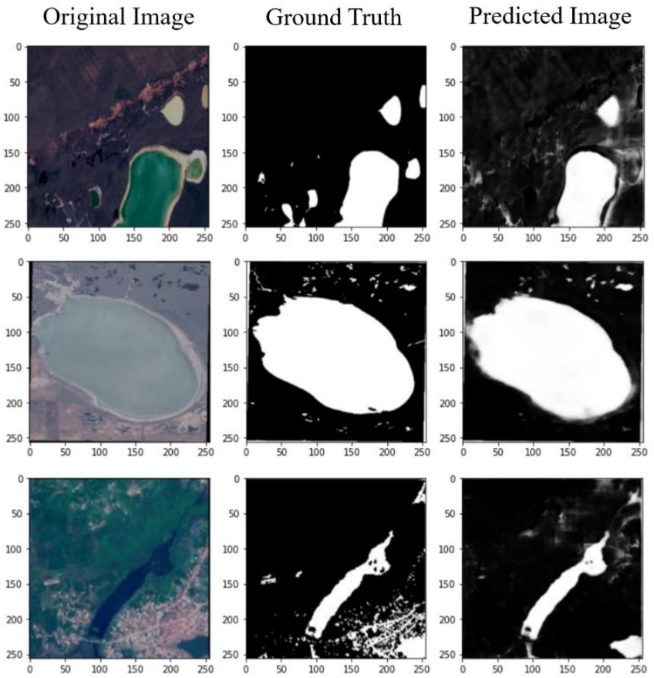

Fig: Above images refer to the results generated by UNet Model

c) Gradio

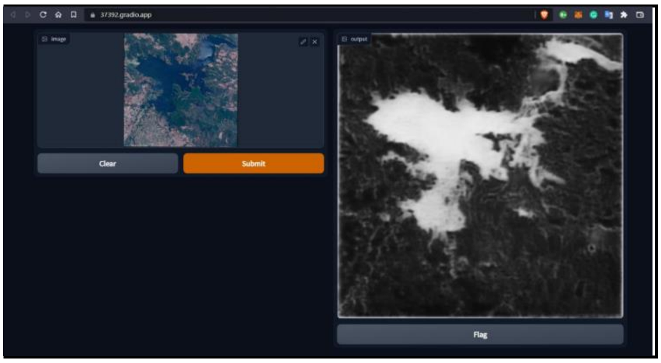

Fig: Gradio application used for Testing Arbitrary satellite image with mask generated on Right Box

CONCLUSION

Extraction of River networks tends to resolve unnoticed issues faced by agricultural professionals or urban territorial planners. The availability of data is a vital part which consists of accessing high-resolution satellite images from various authentic sources. Minor challenges encountered while accessing these satellite images can be described as irregularity of image size, the existence of non-required objects within the image, and noisy images consisting of shadows and clouds. As a result, the selection of set guideline images collection is an exhaustive task but has to be performed in order to train the model and yield accurate prediction results. Hence, in this paper, image segmentation task has been composed and has been trained oa deep learning model to extract water bodies from satellite images. Presently, abundant techniques are developed for image processing but selecting any specific method and applying on scarce dataset will not achieve the goal. Thus, by considering all probable states of images including binary time zone images, noisy images, etc. Image segmentation using global thresholding method and the developed UNet model predicts accurate masks from satellite images consisting of water bodies with a decent dice score of 80.98 %. The future scope includes development of ensemble method in which specific parameters from selected algorithms will be integrated to get better results with observable clear predictions. And secondary scope includes training of developed ensemble model on human nerve system to extract nerves as for a medical or biological application.

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. "LargeScale 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

[12] Liu, Z., Chen, B., & Zhang, A. (2020). Building segmentation from satellite imagery using U-Net with ResNet encoder. 2020 5th International Conference on Mechanical, Control and Computer Engineering (ICMCCE). doi:10.1109/icmcce51767.2020.00431

[13] https://www.kaggle.com/datasets/franciscoescobar/satellite-images-of-water-bodies

[14] https://developers.google.com/earth-engine/datasets/catalog/sentinel

[15] https://earth.google.com/web/@23.38968689,73.97390015,150.00634844a,59857.8614097d,35y,355.85248184h,0t,0r