Deep Learning Based Water Feature Mapping Using Sentinel- 2 Satellite Image

Problem Statement:

As a foremost resource, water has a wide-ranging effect on numerous economic activities. In farming, water manages the scattering of crops. Urban water is significant for the urban ecological community. Precise and systematic location of urban water with remote detecting information is major for urban administration and arranging. Extraction of water information has become a foremost application of Remote Sensing. Mapping of surface water is helpful in a variety of remote sensing applications. For example, estimating the availability of water, and predicting droughts and floods. This study aims to extract various earth objects from medium resolution satellite imagery using deep learning approaches. In general, the proposed technique is relevant to precise water planning with improved spatial goals and exactness, which conceivably encourages water planning and its related investigations and applications on developing Sentinel-2 satellite images.

Background

Water is an important asset for a property scheme on earth. It contributes considerably to the parity of biological systems, the upkeep of global environmental change and in this manner the carbon cycle. The arrangement, extension, shrinkage and vanishing of surface water are imperative elements impacting the environment and local atmosphere changes. Water is additionally a crucial think about socioeconomic development, as a result of it affects several farming, natural and environmental issues after some time. Additionally, the speedy and correct extraction of water asset data will offer fundamental knowledge, that is of nice decent centrality for water asset examination, flood observance, soil protection and disaster[1]. Moreover, Urban water may be some important portion of the urban scheme furthermore plays a crucial job in human life and concrete economic development, like facility, control, tourism, and concrete. Urban water varieties embrace streams coursing through urban regions, and common or synthetic lakes, ponds, and supplies. Nonetheless, in late decades, on account of the effect of human exercises and worldwide environmental change, land use/spread in urban territories has experienced serious changes, prompting emotional changes. This not solely impedes the property improvement of urban water assets, however additionally influences the conventional operate of the urban biological systems. Accordingly, associate objective and correct comprehension of the spatio-fleeting appropriation attributes of urban water is important for urban planning and advancement [2].

Additionally, Sentinel-2 mission, was propelled by European Space Agency (ESA) in 2015, gives another gracefully of overall lined, plainly available optical remote detecting. Contrasted with comparable Landsat arrangement, Sentinel-2 conveys pictures with higher spacial goals, extra otherworldly groups, less return time, and more extensive area, accordingly has indicated decent possibilities in an exceedingly wide fluctuate of fields, similar to land cowl what's more, land use, sea-going science, agrarian applications, and soil analysis. Land surface water location has been one among the premier fundamental regardless imperative remote detecting applications. The water recognition permits the investigation of human changes to the environmental factors in an exceedingly. Water detection is additionally usually one among the essential strides in planning land-use and land-spread, foreseeing waterborne irresistible sickness, overseeing flood peril, evaluating water inadequacy, and surveying water quality. Hence, land surface water recognition extensively adds to remote detecting studies and applications[3]. The European Space Agency's (ESA) Sentinels have begun AN Earth Observation (EO) new period particularly in disaster observance and crisis the executives. Sentinel-2 (S2) crucial progressed spatial goals and recurrence of acquisitions contrasted with past satellite missions creating on the market detached optical knowledge, e.g. Landsat. Landsat knowledge, however, still represent the sole historical consistent archive of EO knowledge for semi permanent flood planning and observance [4].

In certain districts, factors like global environmental change, farming enlargement, increase, and expanding urbanization include caused an ascent inside the interest for water, bringing about uneven characters in give and request, Irrigated agribusiness is one among the most wellsprings of water request, representing over seventieth of the water withdrawal inside the world. The Food and Agriculture Organization of the UN (FAO) assesses that a region of every 324 million hectares is provided with water system foundation inside the world, and eighty fifth. The expanding utilization of water for horticultural water system may cause issues on a local scale [5]. Deep learning is that the learning method that recreates the human mind. It will precisely separate elevated level choices from low-level choices of the info picture [6]. In this paper, a completely unique deep-learning architecture is proposed to extract various earth objects from medium resolution satellite imagery using deep learning approaches. Satellite images analysis is a very important area for various societal needs. Extraction of various forms of objects from satellite image becomes difficult task when the spatial resolution of satellite image is coarse. So, there's a desire to develop an approach which can be used for such satellite image resolution. This can be utilized for urban planning and its management.

Related Work

Chen et al. [1], 2018 , "Urban water bodies extraction" the information for this investigation depend on two classifications of Chinese high-spatial goals remote-detecting pictures for urban water extraction. Which were ZY-3 and GF-2 multispectral pictures. It has been viewed as the complex urban water organizes fluctuation in China. The choosen district for this examination was found in Beijing, Tianjin and Chengdu. This examination has analyzed four remote-detecting multispectral pictures picked up from the ZY-3 and Gaofeng-2 satellites having not at all like area sizes. What's more, the examination region covers three urban region which are encircled by rural water bodies, for example, lakes, lakes, thin waterways and oceanic parks.
This examination says that the tried pictures have been adjusted for radiometric and geometric remedy. Reference water planning is physically digitized by an optical clarification methodology of the high-goals symbolism concerning Google Earth. Isikdogan et al. [2], 2017, "Surface Water Mapping" this examination is an endeavour to introduce the convolutional neural system design that is equipped for learning land spread highlights at numerous scales. Also, it has key modification that adjust their model to the choosen application, along with lower number of trainable boundaries, the examination at a bigger number of scales, and the manner in which the layers are coupled. In addition to this design, they prepared a profound learning based surface water model for Landsat pictures. The proposed model remembers the highlights of water bodies for factors over the globe. These highlights assists with recognizing water from day off, cloud, and territory shadows. This model is straight forward to execute and is quick in application. The problem with this model is that they used Landsat 7 images which has a spatial resolution of minimum 15m but our data is taken from Sentinel 2 with spatial resolution of 10m. Saraiva et al. [3], 2020, "Mapping of irrigation systems" this examination predominantly focuses on programmed recognition and the planning of the area, shape, and zone of focus rotate water system frameworks. The fundamental goal of this examination is to introduce a completely mechanized method which is fast, progressively exact, and increasingly adaptable when contrasted with planning

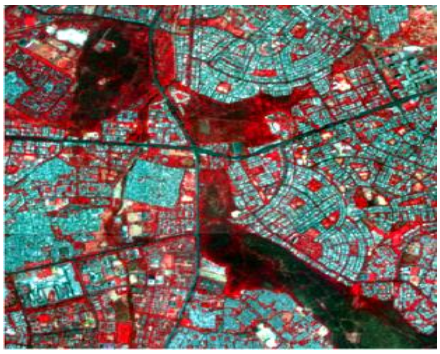
by visual examination. The information for this examination is situated between the conditions of Goiás (GO) and Minas Gerais (MG), in Brazil. Both the picked lattices for this investigation are in the Cerrado biome. Goiás is the state with the second massive focus rotate watered region in Brazil. Wang et al. [4], 2020, “Water identifying using CNN” this examination is an attempt to utilize convolutional neural network (CNN) to extract water bodies from GF-1 pictures. They likewise acquired the possibility of DenseNet and added the up-sampling activity to make a completely convolutional neural system. All the while, the skip layer association was connected in the up-sampling and down-testing activities to better the effectiveness of highlight application. What's more, the examination contrasts this model and the two division systems of SegNet and DeepLab v3+. Just as the investigation analyzes two element extraction systems of ResNet and VGG. Also, besides, the investigation separates the conventional water record practice to perceive their efficiencies in water body acknowledgment.

Wang et al. [5], 2020, “Urban water extraction using deep learning and GEE” this study is an attempt to consolidate two models, in particular GEE with a CNN model, to decide on urban water for another procedure. The studied model of CNN, which is often know as the multiscale convolutional neural system (MSCNN). Z. Wang et al. [6], 2018, “MuWI for accurate water mapping” this concentrate chiefly expects to advance the Sentinel-2 intrinsic multi-spectral water index (MuWI). Furthermore, it produces 10 m water planning without band sharpening. Additionally, the examination intends to separate the precision of various water records in ordering various sorts of water on Sentinel-2. Also, besides, the examination expects to imply an objective, activity-clear evolution of index procedure for binary categorization while making use of OSH (optimal support hyperplane) in SVM (support vector machine). This model was prepared disconnected utilizing chose Landsat pictures and the corresponding water masks. Besides, the structures of the prepared model were transferred to GEE. Also, these systems were utilized to mimic the action of MSCNN's discovery of urban water. Completion of the urban water extraction of Landsat pictures on GEE was performed. Furthermore, the introduced system can be encapsulated as “offline training and online prediction” (OTOP).

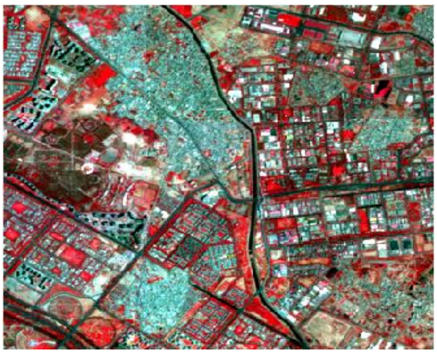
Study Area and Data Resources

Study Area

The study area is situated in India, Delhi NCR, in North and East coordinates. The dataset utilized is Sentinel-2 images from ESA (European space agency) for this study. The height and width of the dataset utilized from Sentinel-2 images is 3455 × 5063 pixels. The Coordinates (In Extent) are 701320.0, 3149340.0:751950.0, 3183890.0 meters.



(a)



(b)

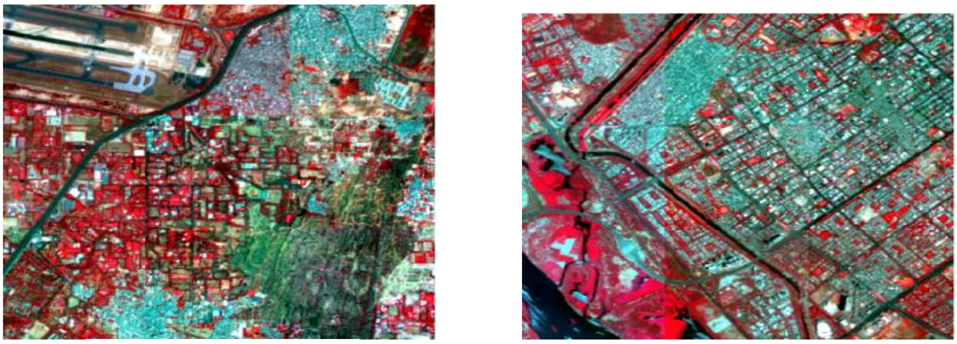


Figure-1. Sentinel-2
RGB image

Dataset

Sentinel 2 Bands and Combinations

Sentinel 2 has 13 groupings in altogether. Multispectral Imager (MSI) sensor in Sentinel 2 transmits 13 groups. The pixel size of these range from 10 to 60 meters. Its channels blue, green, red and near infrared have resolutions of 10 meters. NIR, red edge and short-wave infrared have resolutions of 20m. Cirrus band and aerosol have a resolution of 60 meters. Through band has pixel size of 10, 20 or 60 meters. Sentinel 2 features 2 satellites. First came, a 2015 propelled Sentinel 2A.

Table-1. Overview of Sentinel 2 Dataset

Band	Resolution	Central Wavelength	Description
B4	10 m	665 nm	Red
B3	10 m	560 nm	Green
B8	10 m	842 nm	Visible and Near Infrared (VNIR)

Next came in 2017, Sentinel 2b. Is conveyed by Sentinel-2. Otherworldly groups ranging from pixel size of 10 to 60 meter. Next, their red edge (B5), near infrared NIR (B6, B7 and B8A) and short-wave infrared SWIR (B11 and B12) have 20 meters of ground sampling distance. The application of shading infrared band is intended to highlight solid and unfortunate vegetation. By utilizing the near-infrared (B8) band, Chlorophyll reflecting is especially appropriate. This is the reason why, denser vegetation is red in a shading infrared picture. Urban territories are white, in any case.

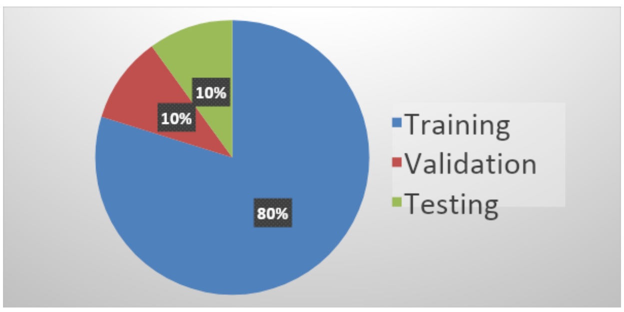


Figure-2. Pie chart of dataset distribution into training , validation and testing.

Methodology

This is our first model. According to the figure 2, we did 2 convolutions in each layer but we did with zero padding. We did copy the values from previous layers using concatenate function but we didn't crop the image so the dimensions are same. We deleted few layers to fit it in our model correctly.

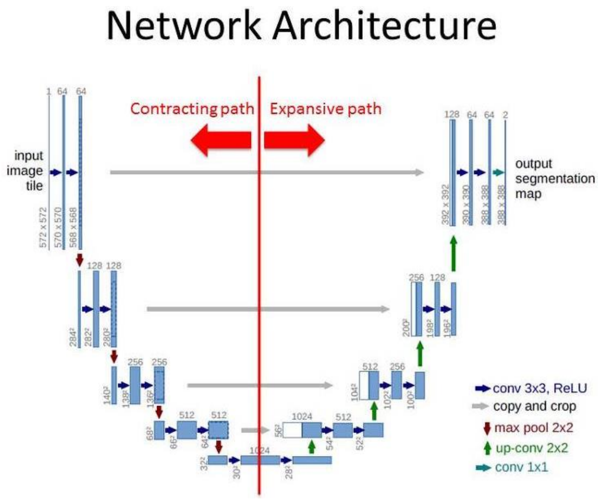


Figure-3. U-net architecture, Ronneberger et al. [7], 2015

This is our second model. We rolled out custom improvements to our model by utilizing encoder-decoder structure. Moreover, we added data augmentation by taking transpose of existing images. This doubled the quantity of images to make it increasingly appropriate for our application. It was evaluated that not many layers were needed as we have just two classes to foresee i.e., waterways and Non-waterways. We didn't use dropout.

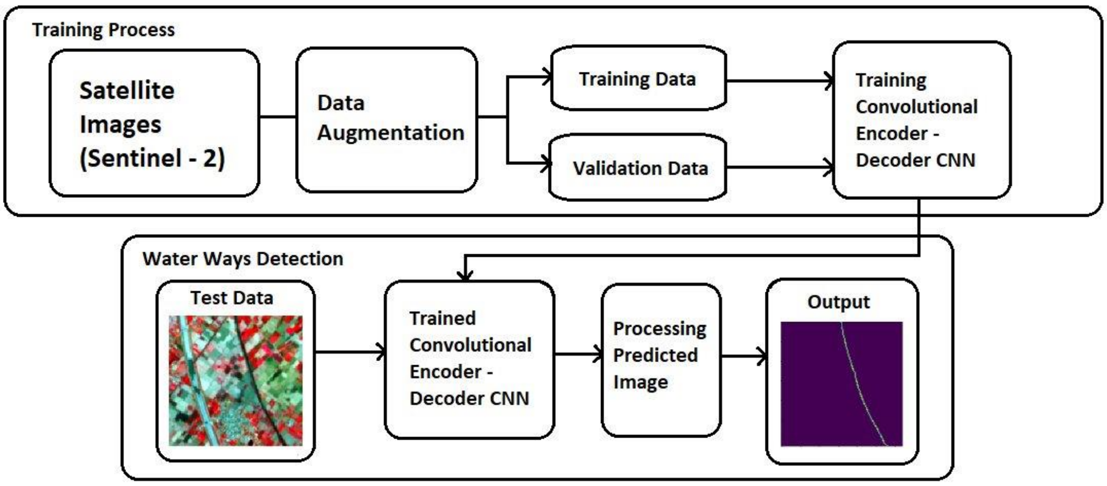


Figure-4. Flowchart of our model

We used min-pooling as it is better than max-pooling for recognizing dark pixels, which is our case. We included Batch Normalization to accelerate the convergence. The ReLu activation function was integrated for each layer except the last one. The Sigmoid Activation function was implemented for the last layer so that we get values in the range of 0 to 1. In the wake of foreseeing the qualities we utilize a threshold value so all the pixels in the predicted image turn out to be either 0 or 1. The Model we presented uses TensorFlow framework and its resources. About 7.17 GB of NVIDIA Tesla K80 GPU and 3.09 GB of RAM takes around 60 minutes to train the model. The equation of computation excluding Minpooling and Batch Normalisation is as follow:

$$f(x) = \mathrm{Relu}\ (\mathrm{Relu}\ (\text{input image} * \mathrm{Conv1}) * \mathrm{Conv2}) \quad (1)$$

Here, 'x' represents the 1st layer.

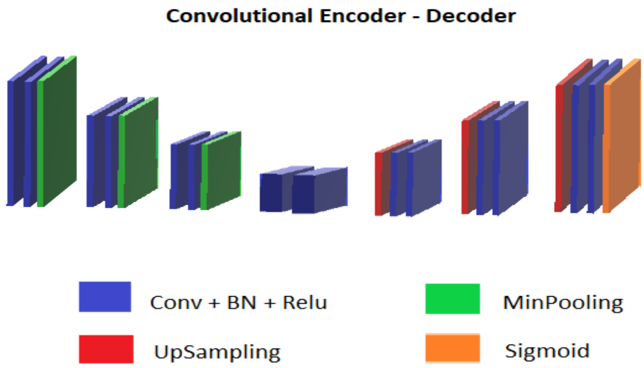


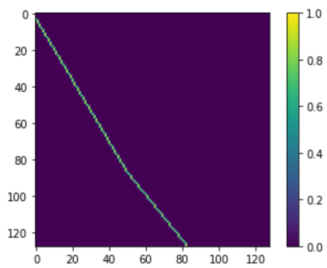
Figure-4. Architecture of the model, Yasrab et al. [8], 2017

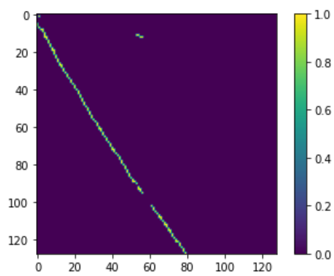
Experimental Design

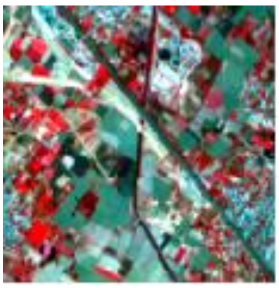
Our first model was not accurately predicting the water ways from a given satellite image and the results are as follows.

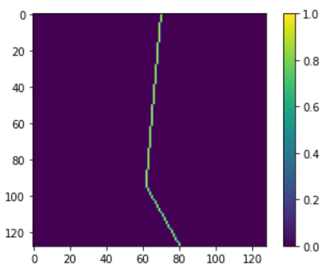
Input image | **Ground Truth** | **Predicted Image**


(a)


(b)


(c)


(d)


(e)

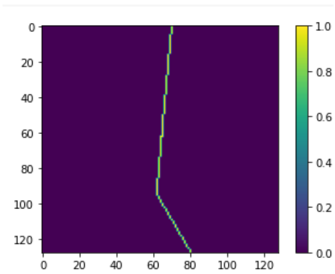

(f)

Figure-5. Result of the water feature mapping using the U-net model. Representation of Actual Satellite image, Ground truth image and predicted output image. (a) and (d) are the input images, and (b) and (e) are the ground truth images, and (c) and (f) are representing the predicted image.

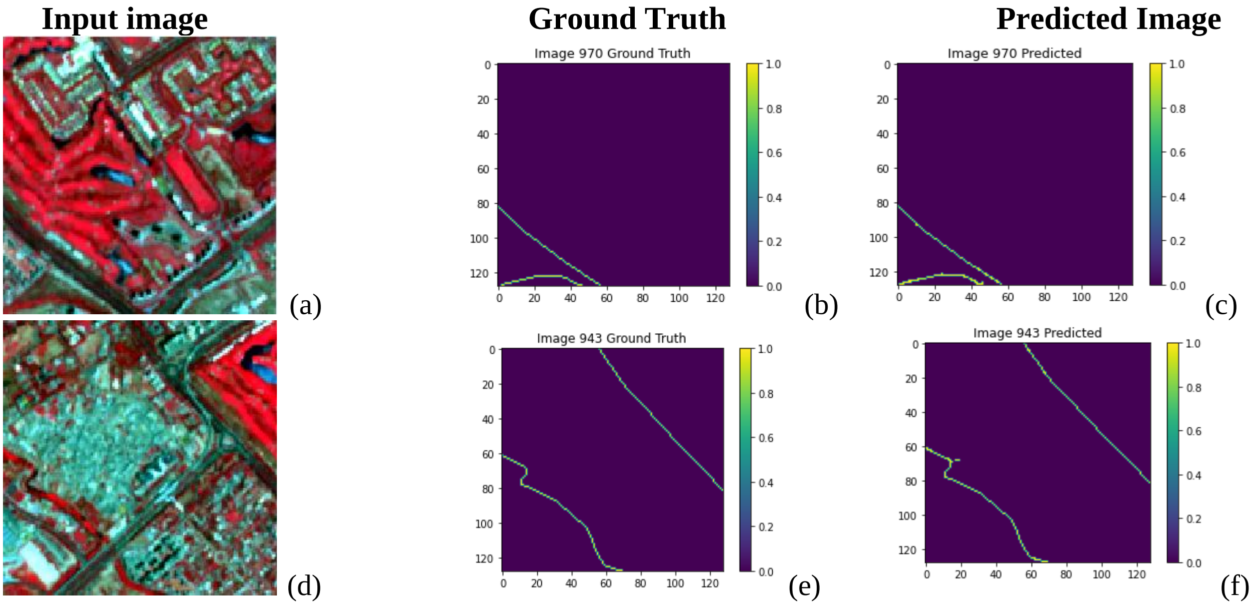


Figure-6. Result of the water feature mapping using our Encoder- Decoder model. Representation of Actual Satellite image, Ground truth image and predicted output image. (a) and (d) are the input images, and (b) and (e) are the ground truth images, and (c) and (f) are representing the predicted image.

Our second model is successfully predicting the water ways with high accuracy from a given satellite image and the results are as follows. The value of 1 implies that there is a water way and 0 implies that there isn't.

Evaluation Measures

The performance evaluation of our model is executed using these measures: Precision, Recall and F1-Score.

Table 2. Validation of results of our model is shown in the table

	Precision	Recall	F1-Score
Modified U-Net architecture	82.01	83.40	82.70
Our Encoder - Decoder architecture	88.11	90.75	89.41

Precision - Precision is the proportion of accurately predicted positive perceptions to the total predicted positive perceptions.

$$Precision = TP/TP+FP \qquad (2)$$

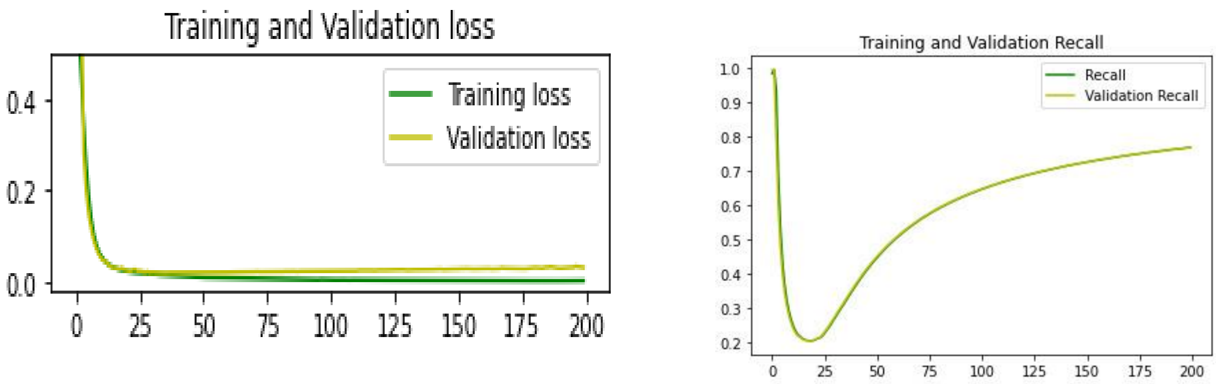
Recall - Recall is the ratio of effectively predicted positive perceptions to the all perceptions in actual class.

$$Recall = TP/TP+FN \quad (3)$$

F1 score - F1 Score is the weighted average of Precision and Recall.

$$F1\ Score = 2*(Recall * Precision) / (Recall + Precision) \quad (4)$$

Here, the 'TP' stands for True Positives,
'TN' stands for True Negatives,
'FP' stands for False Positives,
'FN' stands for False Negatives.



(a) (b)

Figure-7. Representation of Training and Validation loss and recall. (a) Training and Validation loss. The green line represents the Training loss and the yellow line represents the Validation loss. (b) Training and Validation recall. The green line represents the Training recall and the yellow line represents the Validation recall.

1. Conclusion

Urban water is significant for the urban ecological network. We created a Convolutional Encoder – Decoder neural network that predicts the water ways from low spatial goals pictures of Sentinel 2. The model can separate low level features identified with water with precision of 88.11%. The model can be used to extract features other than waterways such as agricultural areas, urban establishments, roadways etc. The neural network model can be integrated with satellite images obtained on a daily basis to see changing water levels, water flow, drought and other important indicators. The limitations of the project is that the model works best only for 2 class segmentation. Incase more classes are to be added, the model will require a good amount of change to maintain accuracy.

Software and Hardware Requirements

Python based Deep Learning libraries have been exploited for the development and experimentation of the project. Tools such as Anaconda Python, and libraries such as TensorFlow, and Keras have been utilized for this process. Training was conducted on NVIDIA GPUs for training the deep learning models for waterways detection from medium resolution satellite images.

References

[1] Y. Chen, R. Fan, X. Yang, J. Wang, and A. Latif, "Extraction of urban water bodies from high-resolution remote-sensing imagery using deep learning," *Water (Switzerland)*, vol. 10, no. 5, 2018, doi: 10.3390/w10050585.

[2] F. Isikdogan, A. C. Bovik, and P. Passalacqua, "Surface water mapping by deep learning," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 10, no. 11, pp. 4909–4918, 2017, doi: 10.1109/JSTARS.2017.2735443.

[3] M. Saraiva, É. Protas, M. Salgado, and C. Souza, "Automatic mapping of center pivot irrigation systems from satellite images using deep learning," *Remote Sens.*, vol. 12, no. 3, pp. 1–14, 2020, doi: 10.3390/rs12030558.

[4] G. Wang, M. Wu, X. Wei, and H. Song, "Water identification from high-resolution remote sensing images based on multidimensional densely connected convolutional neural networks," *Remote Sens.*, vol. 12, no. 5, 2020, doi: 10.3390/rs12050795.

[5] Y. Wang, Z. Li, C. Zeng, G. S. Xia, and H. Shen, "An Urban Water Extraction Method Combining Deep Learning and Google Earth Engine," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 13, pp. 768–781, 2020, doi: 10.1109/JSTARS.2020.2971783.

[6] Z. Wang, J. Liu, J. Li, and D. D. Zhang, "Multi-SpectralWater Index (MuWI): A Native 10-m Multi-SpectralWater Index for accuratewater mapping on sentinel-2," *Remote Sens.*, vol. 10, no. 10, pp. 1–21, 2018, doi: 10.3390/rs10101643.

[7] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 9351, pp. 234–241, 2015, doi: 10.1007/978-3-319-24574-4_28.

[8] R. Yasrab, N. Gu, and X. Zhang, "An encoder-decoder based Convolution Neural Network (CNN) for future Advanced Driver Assistance System (ADAS)," *Appl. Sci.*, vol. 7, no. 4, 2017, doi: 10.3390/app7040312.

[9] Reed, B.; Brown, J.; Vanderzee, D.; Loveland, T.; Merchant, J.; Ohlen, D. Measuring phenological variability from satellite imagery. *J. Veg. Sci.* 1994, *5*, 703–714.

[10] Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems;* Curran Associates: North Miami Beach, FL, USA, 2012; pp. 1097–1105.

[11] Couprie, C.; Farabet, C.; LeCun, Y.; Najman, L. Indoor Semantic Segmentation using depth information. In Proceedings of the International Conference on Learning Representation, Scottsdale, Arizona, 2–4 May 2013.

[12] Hu, F.; Xia, G.S.; Hu, J.; Zhang, L. Transferring deep convolutional neural networks for the scene classification of high-resolution remote sensing imagery. *Remote Sens*. 2015, 7, 14680–14707.

[13] Wang, J.; Song, J.; Chen, M.; Yang, Z. Road network extraction: A neural-dynamic framework based on deep learning and a finite state machine. Int. J. *Remote Sens*. 2015, 36, 3144–3169.

[14] Matikainen, L.; Karila, K. Segment-based land cover mapping of a suburban area—Comparison of high-resolution remotely sensed datasets using classification trees and test field points. *Remote Sens.* 2011, 3, 1777–1804.