AUTOMATED EXTRACTION OF WATER BODIES FROM ALOS-2 IMAGES USING U-Net and ROUGH TRAINING SET

Vaibhav Katiyar (1), Masahiko Nagai(1)

[1] Yamaguchi University, 2-16-1, Tokiwadai, Ube, Yamaguchi, 755-8611, Japan
Email: i502wd@yamaguchi-u.ac.jp, nagaim@yamaguchi-u.ac.jp

KEYWORDS: Convolution Neural Network (CNN), U-Net, Synthetic Aperture Radar (SAR), ALOS-2

ABSTRACT: In this article, authors have suggested extracting water bodies by using one type of Fully Connected Network (FCN) i.e. U-Net over high-resolution Synthetic Aperture Radar (SAR) images from ALOS-2. Due to the limitation of the training samples and binary classification i.e. water or no-water, this study has selected U-Net which gave better performance even with lesser training samples. Using rough training set to extract the water bodies is really a big challenge as SAR images already have noises such as Speckle Noise or Gaussian Noise. However, the study shows that just by using rough training set on the backscatter image (dB) gives very good result in extracting water bodies and even able to differentiate in some of the difficult areas such as extracting water bodies from golf courses which visibly looks similar to water areas in the SAR image.

INTRODUCTION

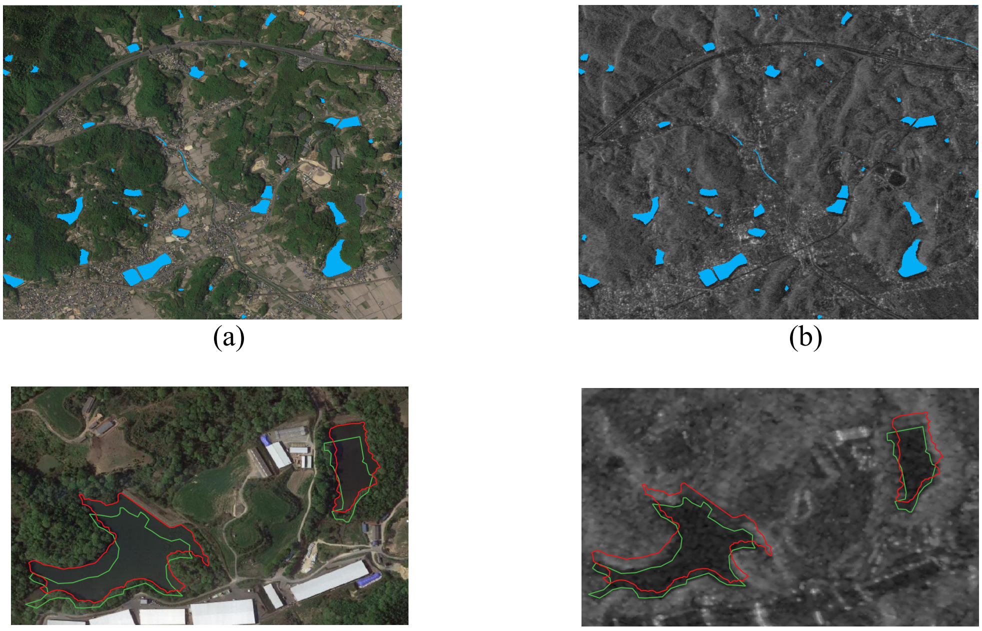
Automated extraction of water bodies in the geospatial domain has various applications such as urban planning, hazard mapping, change detection, etc. This has been specifically highlighted in Sustainable Development Goals 6 (SDG-6) i.e 'Ensure availability and sustainable management of water and sanitation for all". In the sub-goal 6.6 it has mentioned as 'by 2020, protect and restore water-related ecosystems, including mountains, forests, wetlands, rivers, aquifers and lakes' (UNSKP, 2015). Implementation of these objectives is required frequent monitoring of these sites to avoid any encroachment and proper maintenance. However, regular manual survey and monitoring are not practical as it will be more costly and human-resource intensive. Due to these reasons along with the improvement in the technology such as satellite imagery with higher special, spectral and temporal resolution and development of the machine learning techniques- an automated system of water bodies extraction has been propounded. Though most of them are mainly focusing on using optical images (Huang, 2018) such as low & medium spatial resolution satellite images (Mishra, 2015; Jawak, 2015) and High & very high-resolution satellite images (Liwei, 2019; Feng, 2019). While they have produced some very good results but still the limitation of the optical image is always present that is affected by cloud and weather condition. Due to this reason Huang, 2018 have recommended using SAR images along with optical images which can help not only in penetrating the cloud but also the vegetation. Chandran, 2018 is using sentinel-1 (SAR) satellite for extraction of the water-bodies, however, training a separate network for finding the shadow to remove it from the final detected water bodies.

This study is trying to make water bodies detection simpler and more ubiquitous by using simple network architecture and using rough or less accurate training data.

TRAINING DATA PREPARATION

As there was no training data available for ALOS-2 satellite images so training data for water body has been prepared in-house. Single polarized ALOS-2 image (HH) (auig, 2019) of the month of April has been chosen which has a resolution of 10 m with observation width 70 km. Two Zones which has been observed as having more number of water bodies has been selected with the dimension of 15 x 20 Km. For creating the training data faster, the study had used google earth high-resolution optical images with the QGIS digitization tool. The output of this step is saved as a vector file containing polygons of the shape of water bodies. Later on, this vector layer has been converted as raster layer with binary value, 1 for water-bodies and 0 for others, converting it as a mask layer. Even though ALOS-2 PALSAR-2 L1.1 data has been orthorectified using ESA's SNAP (Sentinel Application Platform) toolbox, It was still not very well overlapped, one of the reasons was the usage of 30-Meter (low resolution) SRTM DEM (Shuttle Radar Topography Mission, Digital Evaluation Model) as well as certain limitation in horizontal positional accuracy of google earth (Goudarzi M.A.,2017). Along with it, Mountainous terrain and SAR side-looking geometry were making things more complex for perfect overlapping. So this is why to create training data with better quality, manual shifting of the polygons have to be done, this is a complex task as small water bodies may not be visibly clear in the SAR image. Moreover, the boundary of the water bodies may also be blurred which in turns reduce the quality of training data. Due to this reason, the study utilizes the same polygon which has been drawn on google earth images without shifting them as can been seen in Figure1 (a) and (b). The unevenness of the real ground truth can be seen in Figure 1(c) and (d) where red colour polygon shows the boundary of the water body in google earth and green colour polygon shows the boundary of the same water body in SAR image.

After this, iterative random clipping has been used to get 256x256 size of SAR image for the training set, totalling to 20,000 image tiles. However, high unbalanced data have detrimental effects on the classification or segmentation algorithms(Buda M., 2018). To reduce this problem, the study has used a threshold approach in the final phase of training data preparation i.e. selecting a tile as 'final' only when it has at least 5 % pixels underwater bodies. For calculating this we have used a ratio of 'total number of pixels' with the 'total pixels that belong to water-bodies' from the water-bodies mask layer.



(a) (b)

(c) (d)

Figure 1: Training data preparation. (a) and (c) are the clipped google earth images, blue color in them shows digitized water bodies area and (b) and (d) are the clipped ALOS-2 images after converting into backscatter (sigma zero - db) format using SNAP.

METHODOLOGY

This study has adopted U-Net architecture (Akeret, 2017; Akeret, 2018) with some modifications to handle the satellite images, this network is running on the two Titan-V GPU with 12 GB memory each. The language and supporting libraries used are the Python and it's supporting packages along with the TensorFlow and Keras.

Network and architecture

U-Net is running with the 7 convolution layers in the encoder part and 6 convolutions in the decoder part. Due to the limited number of training samples there was the high possibility of overfitting and to avoid this, certain drop-out layers (regularization) has been inserted with probability (p) = (0.8, 0.5) for the different hidden layers of the network. Leaky ReLU activation function has been used in the hidden layers and sigmoid in the output layer due to binary classes.

Training

Total 20,000 images of size 256x256 have been used for training the network with the validation-split 30 per cent. Which means at one time 14,000 images has been used for training and remaining 6000 has been used for validation purposes dynamically. Just a little over 100 epoch model shows convergence with the batch size 25 (Figure 2).

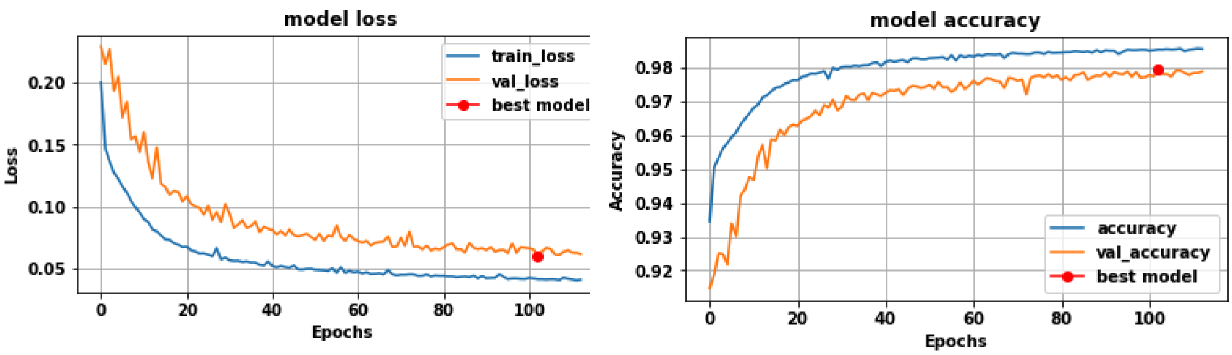


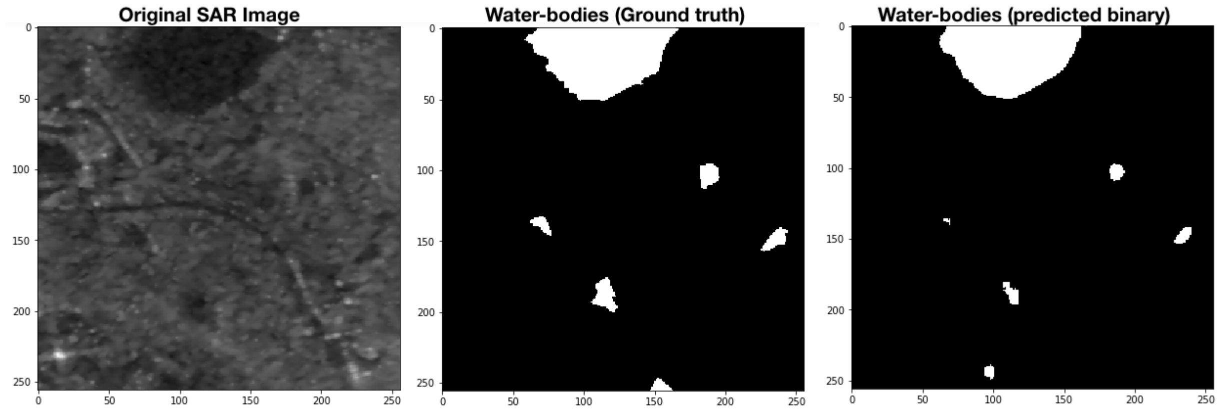
Figure 2: model loss and model accuracy chart are showing that gradually loss is decreasing and accuracy is increasing. Which gives an insight that over-fitting has been avoided.

Testing

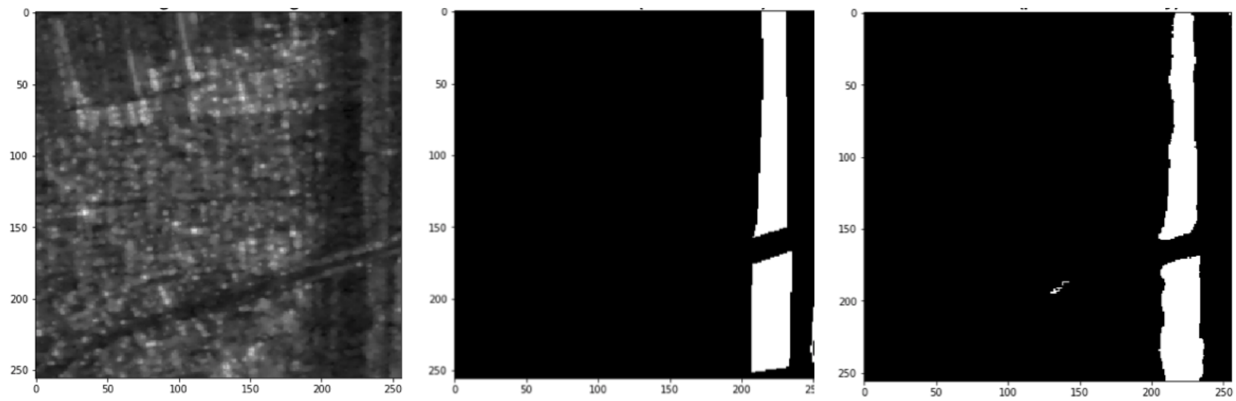
For testing and evaluating the model 1500 images of the same size (256x256) has been used which have not been seen by the model before. This set has been randomly divided into three parts to test the model. The predicted probability image has been converted to the binary prediction using a threshold of greater than 50%. For evaluating the model, Intersection over Union (IoU) matrix has been chosen which can be defined as (Eq1)-

$$IoU = \frac{(Ground\ truth \cap Predicted)}{(Ground\ truth \cup Predicted)} \quad \text{(Eq1)}$$

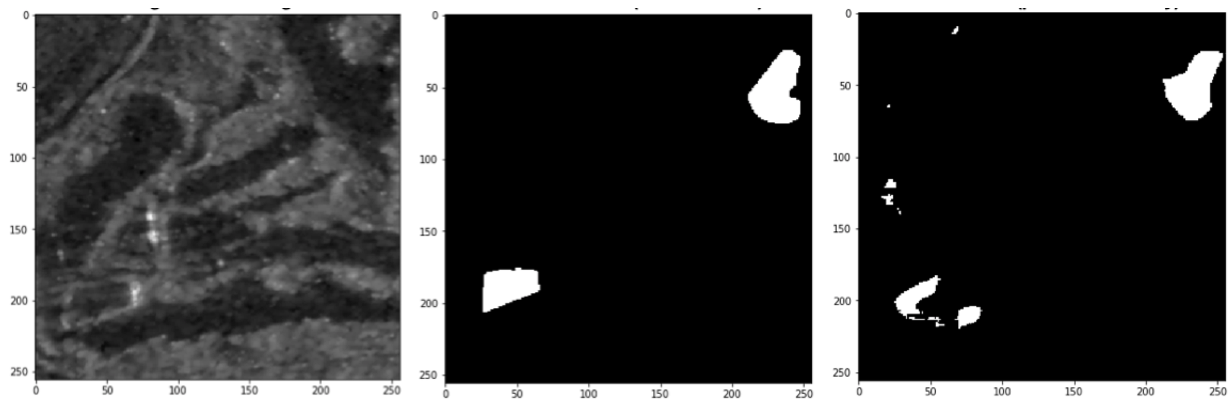
IoU for the parts of the test set has been falling in the interval of 0.69 to 0.76 with the loss hovering in the interval of 0.22 to 0.28. This IoU seems at the lower side, however, we need to realize that IoU is best measured when the bounding box is pretty good. In this study, the ground truth has been rough so even when the model will predict exact object still IoU will not be close to 1.0.



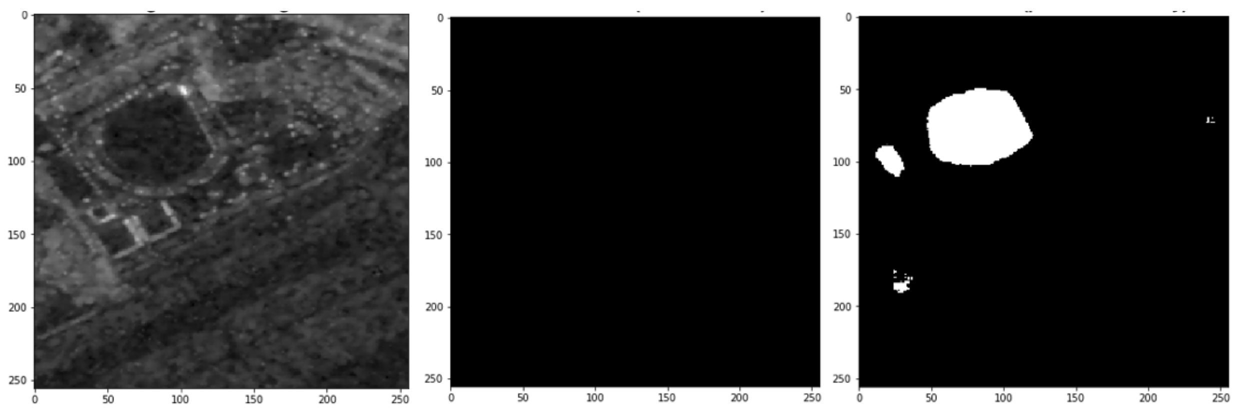
(a) Small water bodies



(b) Rivulet



(c) Water bodies in between golf course



(d) Baseball stadium

Figure 3: Left side image shows SAR image, the middle image shows the ground truth and the right side is showing predicted water bodies in binary form (after thresholding)

RESULT AND DISCUSSION

Some of the predicted output of the images by the trained model has been shown in figure 3. If we see figure 3(a) where all the water bodies even though they are small has been identified the model though the extent of them seems a little bit reduced. In figure 3(b) the rivulet with the overbridge has been captured properly. Figure 3(c) was really a challenging task as small water bodies were in the golf areas which itself looks very similar to the water areas in the SAR image. However, the model has been able to distinguish between water bodies and golf field beautifully. Even the water body at the bottom left corner has not been detected more clearly due to the presence of power lines which create more backscattering. Though the still model has wrongly predicted few areas as water. On the other hand in Figure 3(d) model has predicted the completely wrong result, the baseball field is very flat and it shows a very strong probability to classify it as a water area. While there are other sports fields and with a similar flat surface, the model does not recognise them as water bodies. In the case of baseball field specifically, it shows stronger bias towards the water, it may be a case of shape which model learn due to the presence of many lakes in the training set. Still, this point needed to be explored further in the study.

CONCLUSION

While SAR data is already becoming ubiquitous, and many new SAR satellites are planned to be launched. The availability and accessibility of SAR images will not remain much of a problem. Such kind of models can be very helpful to process SAR images automatically and with less amount of human resource and expenditure. The above discussion shows that a simple model with even a rough training set is able to provide a good accuracy over SAR images. Therefore, the availability of a better annotated and larger dataset can help to better utilize SAR images for many areas including for disasters such as flood monitoring, urban development, water ecosystem management etc.

REFERENCES

References from Journals and Conference proceedings:

Akeret, J., Chang, C., Lucchi, A., & Refregier, A. (2017). Radio frequency interference mitigation using deep convolutional neural networks. Astronomy and computing, 18, 35-39.

Buda, M., Maki, A. and Mazurowski, M.A., (2018). A systematic study of the class imbalance problem in convolutional neural networks. Neural Networks, 106, pp.249-259.

Chandran, B., Remote, N., & Centre, S. (2018). Water body extraction from SAR images, (December), 0–12.

Duan, L., & Hu, X. (2019). Multiscale Refinement Network for Water-Body Segmentation in High-Resolution Satellite Imagery. *IEEE Geoscience and Remote Sensing Letters*, *PP*, 1–5. https://doi.org/10.1109/lgrs.2019.2926412

Feng, W., Sui, H., Huang, W., Xu, C., & An, K. (2019). Water Body Extraction from Very High-Resolution Remote Sensing Imagery Using Deep U-Net and a Superpixel-Based Conditional Random Field Model. *IEEE Geoscience and Remote Sensing Letters*, *16*(4), 618–622. https://doi.org/10.1109/LGRS.2018.2879492

Kaplan, G., & Avdan, U. (2017). Object-based water body extraction model using Sentinel-2 satellite imagery. *European Journal of Remote Sensing*, *50*(1), 137–143. https://doi.org/10.1080/22797254.2017.1297540

Li, L., Yan, Z., Shen, Q., Cheng, G., Gao, L., & Zhang, B. (2019). Water body extraction from very high spatial resolution remote sensing data based on fully convolutional networks. *Remote Sensing*, *11*(10). https://doi.org/10.3390/rs11101162

Meng, L., Zhang, Z., Zhang, W., Ye, J., Wu, C., Chen, D., & Song, C. (2019). An Automatic Extraction Method for Lakes and Reservoirs Using Satellite Images. *IEEE Access*, *7*, 62443–62456. https://doi.org/10.1109/ACCESS.2019.2916148

Mishra, K., & Prasad, P. R. C. (2015). Automatic Extraction of Water Bodies from Landsat Imagery Using Perceptron Model. *Journal of Computational Environmental Sciences*, *2015*, 1–9. https://doi.org/10.1155/2015/903465

Mohammad Ali Goudarzi & René Jr. Landry (2017) Assessing horizontal positional accuracy of Google Earth imagery in the city of Montreal, Canada, Geodesy and Cartography, 43:2, 56-65, DOI: 10.3846/20296991.2017.1330767

O. Ronneberger, P. Fischer, and T. Brox, (2015). “U-net: Convolutional net- works for biomedical image segmentation,” in International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, 2015, pp. 234–241

References from websites:

UNSKP 2015. United Nations Sustainable knowledge platforms. Retrieved 14 Aug 2019. https://sustainabledevelopment.un.org/sdg6

ALOS-2/ ALOS User Interface Gateway. Accessed 17 July 2019. https://auig2.jaxa.jp/openam/UI/Login?

Akeret J., 2018. TensorFlow Unet. Retrieved August 2019, from https://github.com/jakeret/tf_unet