Journal of Advanced Research in Applied Sciences and Engineering Technology

Journal homepage:
https://semarakilmu.com.my/journals/index.php/applied_sciences_eng_tech/index
ISSN: 2462-1943

A River Segmentation for Flood Monitoring with Atrous Convolution via DeepLabv3

Nur Adilah Hamid[1], Mohd Ikmal Fitri Maruzuki[1], Ahmad Puad Ismail[1], Saiful Zaimy Yahaya[1], Fadzil Ahmad[1], Harits Ar Rosyid[2], Zainal Hisham Che Soh[1,*]

[1] Electrical Engineering Studies, College of Engineering, Universiti Teknologi MARA, Cawangan Pulau Pinang, 13500 Permatang Pauh, Pulau Pinang, Malaysia
[2] Jurusan Teknik Elektro, Fakultas Teknik, Universitas Negeri Malang, Kota Malang, Jawa Timur 65145, Indonesia

ABSTRACT

Flood has been identified as a common natural disaster for years. This is the evidence of the effect cause by heavy rainfall which then lead to damages of infrastructure and deaths. The presence of this natural disasters can cause a lot of problems and risk especially to human being. The prevention of flood is almost impossible as it is a natural phenomenon. In this work, we proposed a water segmentation technique to analyses the images of river in term of water at the area from the camera which will automatically detect anomalies such as sudden water increase. The Deep Learning segmentation algorithm DeepLabv3 and DeepLabv3+ are trained and tested for the task of water segmentation and the performances are compared with previous works. In our finding, the accuracy obtained by our proposed method DeepLabv3 is 97.07% thus achieved the state of art in performing the task of water segmentation. Thus, DeepLabv3 model is suit and practical in the solving the flood issue.

Keywords:

Image segmentation; River water segmentation; Fully convolutional networks; DeepLabv3; Flood management

1. Introduction

Floods are a natural disasters phenomenon that happens as the cause of massive flow of water filling up the banks of a river. When storms occur, they have horrible impacts especially on people as this event affect their daily activities and the bad thing is this effect usually last at least a minimum of a week. Flood that happens as the cause of rainfall heavily not only contributes to the damage infrastructure but also deaths and horrible trauma to the victims [1]. In upcoming years, floods are predicted to become even worse with the change of climates. To this regard, the problems will become even more complicated. A storm surge also causes floods to happen in coastal areas. Once a dam collapses if it is triggered, the area of the downstream will overflows. The prevention of floods is almost impossible as it is a natural phenomenon. However, a method in detecting and monitoring the increase of river water can be created so that there will be an early warning thus helping the efficiency process of evacuation and so on. The easiest way to objectively assess whether a warning

* *Corresponding author.*
E-mail address: zainal872@uitm.edu.my

https://doi.org/10.37934/araset.56.2.6273

threshold has exceeded its limit is with the use of static cameras pointing straight to the riverbed then compared with the previous measurements water level. Even so, the monitoring of video cameras in manual way is very costly. In this paper, we propose for the use of water segmentation technique to analyse images from the camera. This technique can immediately track irregularities such as the sudden rise in water. To this regard, deep learning model have been effectively implemented for the task of image segmentation such as image classification etc. The applicability of deep learning model to the flood images in predicting the level of the flood scene in the images have been studied. Thus, we will train, and test deep learning neural network called DeepLabV3 and DeepLabV3+ and the performance will be compared with the other three deep learning algorithms since there is no other method that seems to be accurate in detecting the flood problem. In this paper, the finding results, and the potential use of this tool in addressing early warning in the presence of floods will be discussed.

Fully Convolutional Networks (FCN) was proposed by Long *et al.,* [2]. FCN comes from a combination of connected layers such as convolution, pooling, and up sampling. There are two key differences between FCN and Convolutional Neural Networks (CNN). First, whereas CNN requires fully connected layers in its architecture, FCN requires none. Thus, this reduces the number of learning parameters in FCN. Second, CNN produces class scores as its output. FCN on the other hand, produces structured output or an array of pixelwise values that can be used to identify and predict object localization. FCN has been shown to perform pixelwise prediction in semantic segmentation problem more efficiently than the other segmentation techniques [3,4]. The overall architecture of FCN and its input output is shown in Figure 1.

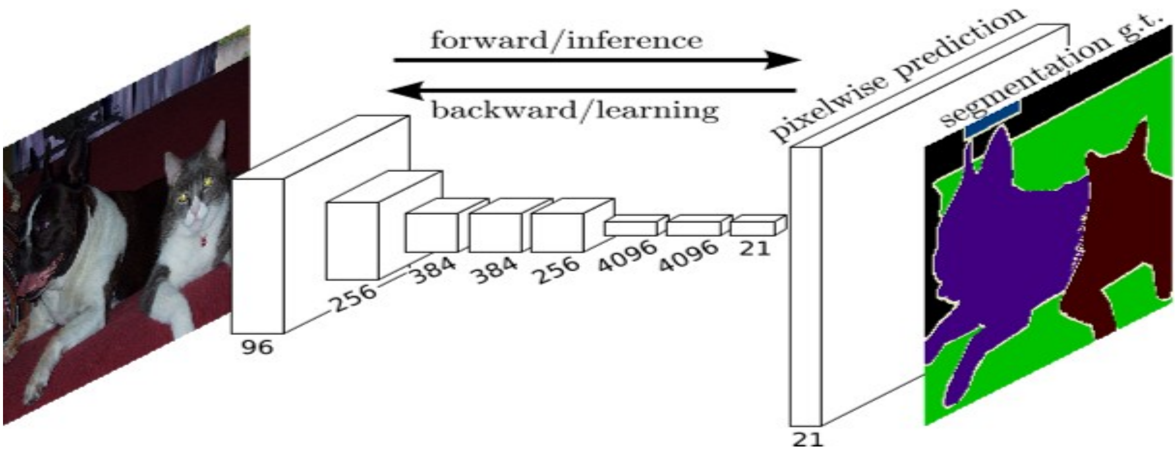


Fig. 1. Overall architecture for fully convolutional networks

The FCN usually consists of several operational layers, where each layer of data in FCN is a three-dimensional array of size $h \times w \times d$. h and w are spatial dimensions, and d is the feature or channel dimension. The first layer is the image, with pixel size h × w, and d colour channels. The subsequent layers in the FCN perform either convolution, pooling, activations, or deconvolution. In the convolution layer, the features in the data are extracted and interpreted by convolving the data with a group of kernels. The spatial output from this operation is reduced from the current convolution layer to the next convolution layer. Thus, this group of processes is called down sampling. Meanwhile, the deconvolution layer calculated the data localization by deconvolving the output from the down sampling layers. The spatial output from operation is increased from current deconvolution layer to the next layer. Thus, this group of process is called up sampling. To further recover the fine-grained spatial information lost in the convolution layers, skip connections are normally used. A skip connection is a connection that bypasses at least one layer. Here, it is often used to transfer local information by concatenating or summing feature maps from the down sampling path with feature maps from the up-sampling path. Merging features from various resolution levels helps combining context information with spatial information.

FCN is shown to perform well in image segmentation. Kai Kang *et al.,* proposed the use of FCN for crowd segmentation where the entire images is taken as inputs and segmentation maps by one forward propagation step is directly output [5]. While Mark *et al.,* have successfully proposed a deep fully convolutional crowd counting model which can perform precisely [6]. There are several attempts to implement FCN in river segmentation problem [7]. However, the produced results from FCN are no so promising.

Pix2Pix is a model designed for general purpose image-to- image translation. Pix2Pix is one example of Generative Adversarial Network, or GAN. The approach was presented by Isola *et al.,* in their 2016 paper titled “Image-to- Image Translation with Conditional Adversarial Networks” and presented at CVPR in 2017 [8].

Figure 2 [9], shows the architecture of GAN model to train text- to- text image. This network consists of two pieces which are the Generator and Discriminator. Their functions are both different. The Generator function in transforming the input image in getting the output image. As for the Discriminator, it is function in predicting the image generated by the generator by measuring the sameness between the input images to an unknown image whether it comes from the target image of the dataset or the generator. Figure 2 also visualizes the sequential processing of the model and the low-resolution images or DCGAN up samples vectors to obtained high- resolution images. As seen on the figure, the layer of the deconvolution increases the spatial resolution image. Not only that, but the depth of the feature maps is also seen to be decreases for each layer. While for discriminator part, the convolutional layers decrease the spatial resolution, and the depth of the feature maps increases.

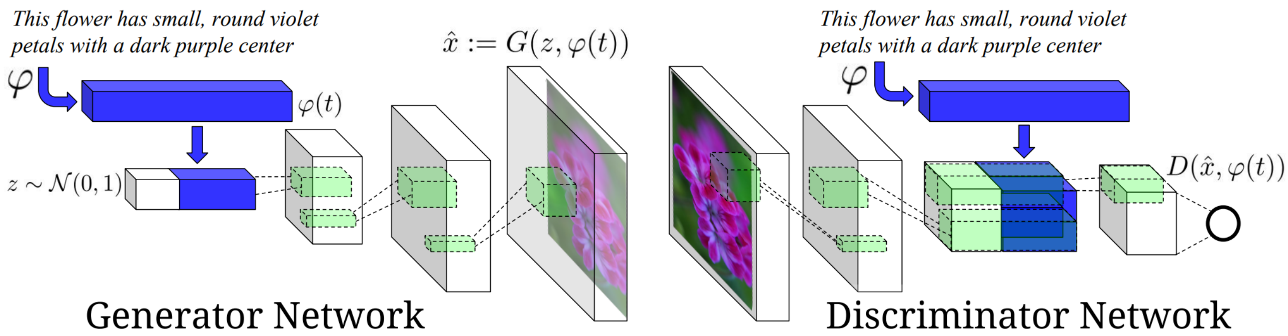


Fig. 2. Overall architecture for Pix2Pix

For generator part, the noise vector z and the text embedding act as input where both are filtered across a fully connected layer so that it can be concatenated with the noise vector. In transforming the input vector to synthetic image, deconvolution layer is used. As for discriminator part, it compressed the text embedding into a 128x1 vector and then reshaped the spatial dimensions into 4x4 matrix. Next, the network then performs depth-wise concatenated with the image representation. The discriminator then takes the image and derived it through convolution layers with BatchNorm and leaky ReLUs after it has been convolved for several times. In this part, before the convolved is over, the embedding is concatenated into the original image matrix. Pix2Pix GAN is shown to perform well in image segmentation. Xian Wu *et al.,* had proposed the implementation of Generative Adversarial Networks (GANs) in image synthesis and editing by training the two main parts which are generator and discriminator in a competitive way [10]. While Lebedev *et al.,* have successfully proposed a modern type of CNN as Conditional Adversarial Networks in change detection in image which results shown a reliable and powerful enough in detecting on both synthetic and actual images [11]. There are several attempts to implement Pix2Pix in river segmentation problem [7]. However, the finding results from Pix2Pix are not really encouraging.

Tiramisu is another solution for semantic segmentation problem. Gao Huang *et al.,* [12] introduced Tiramisu architecture by extending DenseNet, a type of CNN which is built from dense convolutional blocks and pooling operations. Tiramisu characteristic is known for being more precise, efficient, and deeper to train as the connections between layers are closer to the input and output. Tiramisu is composed from a hundred layers of dense blocks and pooling operations. Dense convolutional blocks are iterative concatenation of previous feature maps. Compared to other traditional convolutional networks, Tiramisu have more captivating benefits in terms of encouraging the reuse of feature, reinforce the propagation of feature and reduce the vanishing-gradient problem and number of parameters. Tiramisu has shown excellent results on image classification tasks. Hai *et al.,* had proposed for Fully Convolutional DenseNet with multiscale context for automated breast tumour segmentation where different field of views of image features are acquired without adding number of parameters to avoid over fitting by concatenated multiple sampling rates of Atrous [14] while J´egou *et al.,* had proposed the use of Tiramisu on urban scene benchmark datasets such as CamVid and Gatech and achieved the best results [15]. There are several attempts to implement DenseNets in river segmentation problem [16]. However, the obtained results are still inaccurate.

This paper presents river segmentation for flood monitoring using Atrous Convolution via DeepLabv3. The remainder of the paper is organized as follows. Section 2 describes the methodology of river image segmentation using Atrous Convolution via DeepLabv3 algorithm. Next, in section 3 discussed the experimental result of DeepLabv3 and DeepLabv3+ performances. Finally, section 4 provides the concluding remarks of this work.

2. Methodology

DeepLab is an open-sourced deep convolutional model which known to perform the best semantic segmentation. The model of DeepLab is developed and launch by Google in 2016. Since then, the model has evolved which then produced. DeepLabv1, DeepLabv2, DeepLabv3 and the new one is DeepLabv3+. Due to the improvement of model, their architecture also had improvised. As for DeepLabv3, it is seen to perform fast and accurate for semantic segmentation in real time. Its platform incorporates many important principles for deep learning in computer vision such as spatial pyramid pooling (ASPP) and encoder-decoder architectures. ASPP function in getting information of image at different scales while encoder function to reduce the scale of image to a feature vector and decoder to enlarge the simplified feature vector back to the image dimensions.

Figure 3 shows the architecture of DeepLabv3 [17]. In performing semantic segmentation, this network made a different approach form the designs of encoder-decoder by practicing multi-scale contextual features and controlling signal decimation. DeepLabv3 uses Atrous spatial pyramid pooling (ASPP) [18-20] module which had been improvised with the image-level features and batch normalization. The used of ASPP in this network is to acquire multi-scale data information. It improves the accuracy by compensate various scales of objects by implementing multiple Atrous convolution with various sampling level to the input function map and is fuse together. The results of the analysis are collected by up-sampling. In the ASPP network, at the top of the function map derived from the spine, four parallel Atrous convolutions of different Atrous concentrations are added to the segmentation of the target at different scales.

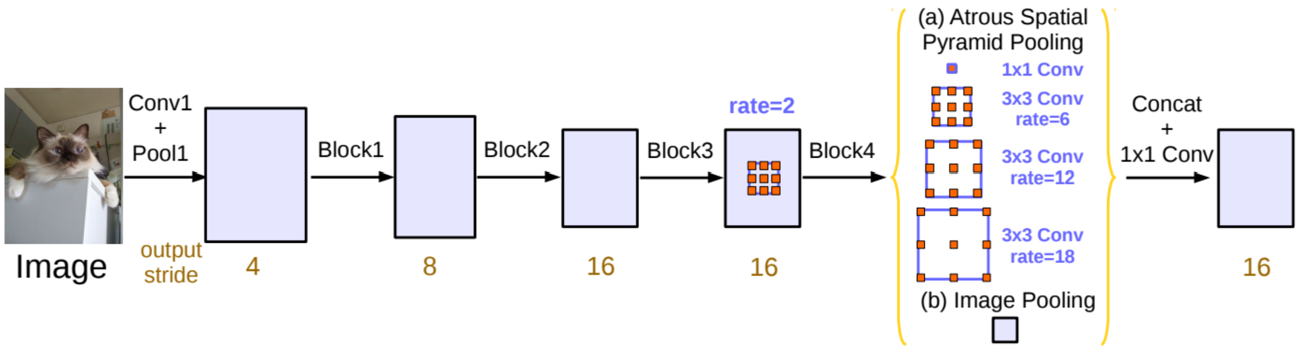


Fig. 3. Overall architecture for DeepLabv3

To add global data information, image-level features are applied. This is done by applying global average pooling on the last feature map of the backbone. Right after performing the operations in parallel, the output result along the channel of each operation is concatenated and 1x1 convolution is implemented to obtain the output. Based on the Figure 3, when the output stride is equal to 8, the rates are double. The output features are then concatenated and passed through a 1x1 convolution along with batch normalization and 256 filters before it generates the logit class output of the final 1x1 convolution.

2.1 Dataset

The dataset consists of a total of 300 images which has been collected through Google, created by self-gathered images and obtained from the surveillance camera in riverbeds. Colour of the water, angle, turbulence and illumination variation are considered during dataset collection. There is a correct ground truth file for each image, consisting of a two-dimensional binary zero matrix for pixels portraying background information, and one for pixels containing water information. Table 1 shows the details of the dataset.

Table 1
The number of pixels and percentage of water dataset [7]

Smallest Image	118x158 pixels
Biggest Image	2448x3264 pixels
Mean Size of Image	550x826 pixels
Min Percentage of Water	6.31%
Max Percentage of Water	91.37%
Mean Percentage of Water	39.57%

2.2 Evaluation Metrics

The proposed deep learning algorithms results is analysed in term of the two main semantic segmentation metrics which are the Mean Intersection over Union (MIoU) and the Pixel-wise accuracy (Pa). As for the dataset, the images are separated into 25% for test and the balance 75% for training. Below is the formula for both MIoU as written in Eq. (1) and the Pa as written in Eq. (2):

$$MIoU = \frac{(1/C)\sum_{i} n_{ii}}{t_i+\sum_{j} n_{ji}-n_{ii}} \quad (1)$$

$$Pa = \frac{\sum_{i} n_{ii}}{\sum_{i} t_{i}} \quad (2)$$

where the n_{ii} represents the pixels from class *i* which have been correctly classified, the n_{ji} corresponds to the number of pixels from class *i* which have been wrongly classified as belonging to class *j*. *C* is the total number of classes and finally t_i is the total number of pixels of class *i*.

3. Results and Discussions

This section discusses the results obtained for river image segmentation using Atrous convolution. Firstly, the benchmarking with existing work in term of quantitative analysis of the proposed and previous works is reported. Secondly, the comparison between DeepLabv3 and DeepLabv3+ in training and validation phase. Finally, the qualitative results of River Segmentation to evaluate the segmentation performance.

3.1 Benchmarking with Existing Work

The quantitative analysis of the proposed and previous works is reported in Table 2 with respect to the statistical measurement of mean and standard deviation along with two semantic metrics. From Table 2, the best results are DeepLabv3 in terms of both MIoU and Pa. Based on the results obtained, our proposed algorithm DeepLabv3 shows the best output results compared to the other algorithms with an accuracy of 97.07% which automatically proves that DeepLabv3 have a higher potential in performing river segmentation tasks. Moreover, DeepLabv3 achieves the lowest standard deviation, which suggests that the results are consistent among the different images. Even DeepLabv3+ achieved an accuracy of 96.32% which put the model in second place in performing segmentation but still this model is better than existing algorithms. The third best is achieved by Tiramisu with an accuracy of 90.47% followed by Pix2Pix with achievement of 84.70% while FCN-8s appeared to perform the worst with the lowest accuracy result which is only 82.80%.

Table 2
Mean Intersection Over Union, MIoU (%) and Pixel-Wise Accuracy, Pa (%) of Existing methods on River Dataset

	MIoU[%]		Pa [%]	
	Mean	Std	Mean	Std
FCN-8s [7]	70.05	14.92	82.80	11.04
Pix2Pix [7]	72.25	14.27	84.70	10.05
Tiramisu [7]	81.91	13.74	90.47	9.16
DeepLabv3+(Ours)	92.02	8.77	96.32	4.79
DeepLabv3 (Ours)	93.52	6.37	97.07	3.40

3.2 Comparison between DeepLabv3 and DeepLabv3+ in Training and Validation

The data and results obtained from dataset during training and validate are evaluated and compared for both model DeepLabv3 and DeepLabv3+. By doing so, the effectiveness of our proposed model in performing image segmentation can be determined. Figure 4 shows the training Pa and training MIoU for both our method DeepLabv3 and DeepLabv3+. As for Pa, both models obtained the same percentage of accuracy of 0.992 while for MIoU, DeepLabv3 achieved 0.9815 and DeepLabv3+ is 0.9832.

During validation, the analysis of the accuracy of DeepLabv3 model for MIoU and pixel accuracy are higher than DeepLabv3+ as shown in Figure 5. MIoU of 0.9225 is achieved while the Pa is 0.9632. To this regard, it can be concluded that the DeepLabv3 model shows more potential to be employed in river segmentation as compared to DeepLabv3+.

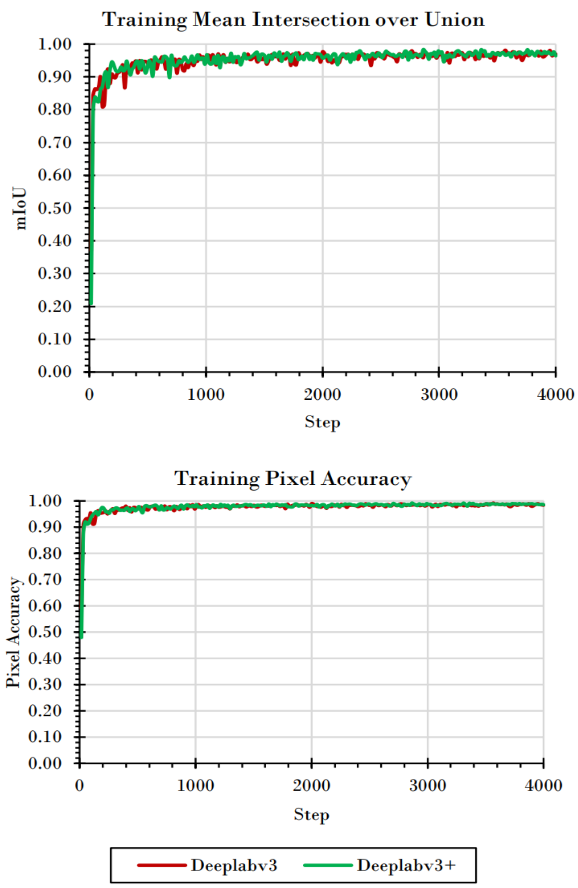


Fig. 4. Graph of Training Mean IoU and training pixel accuracy

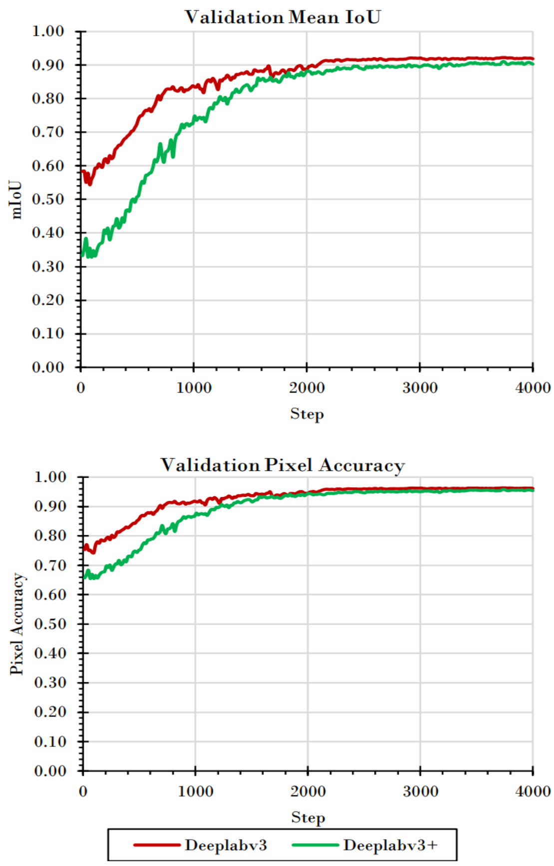
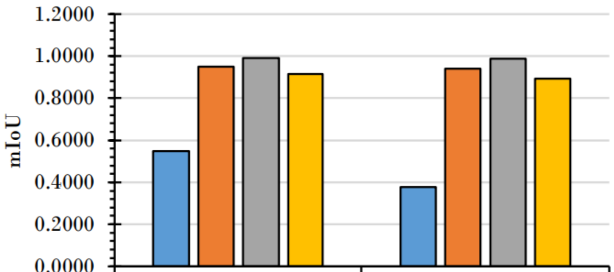


Fig. 5. Graph of validation Mean IoU and validation pixel accuracy

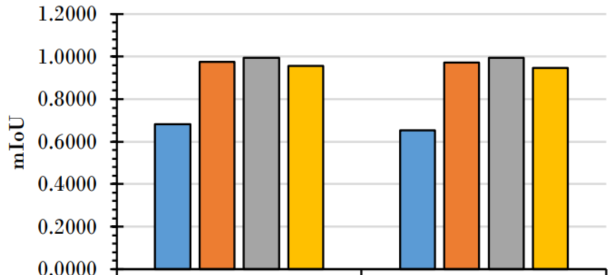
3.3 Comparison between Deeplabv3 and Deeplabv3+ on Different Images

Figures 6, 7, 8, 9 and 10 shown the results obtained for River MIoU, Background MIoU, Overall MIoU, pixel accuracy and the inference time per image for both DeepLabv3 and DeepLabv3+ respectively. The results obtained between these two methods is observed and compared to know which method is more effective in performing segmentation in different images. The mean value shown that DeepLabv3 is performing better in comparison to DeepLabv3+ in all measured items. In River MIoU Mean value, the DeepLabv3 algorithm recorded best output results compared to the DeepLabv3+ algorithms with an accuracy of 91.45%. These also similar for the Background MIoU Mean value, DeepLabv3 also have better result of 95.60% as compared to DeepLabv3+ of 94.71%. In Overall MIoU Mean value depicts that DeepLabv3 perform better with an accuracy of 93.52% as illustrated in Figure 8.



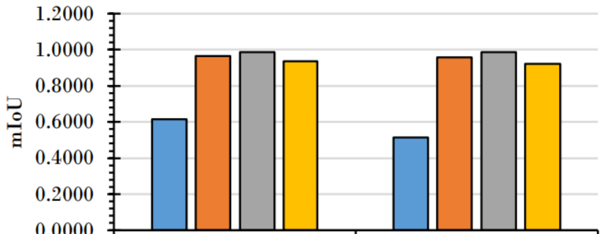
	Deeplabv3	Deeplabv3+
min	0.5474	0.3784
median	0.9497	0.9389
max	0.9904	0.9873
mean	0.9145	0.8933

Fig. 6. River Mean Intersection over Union



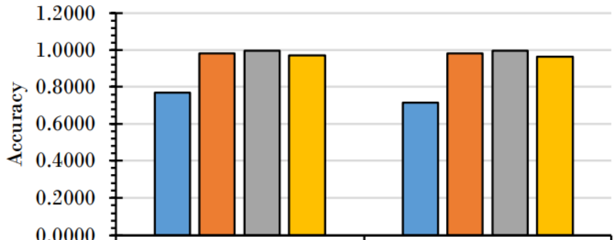
	Deeplabv3	Deeplabv3+
min	0.6803	0.6526
median	0.9744	0.9714
max	0.9931	0.9941
mean	0.9560	0.9471

Fig. 7. Background Mean Intersection over Union



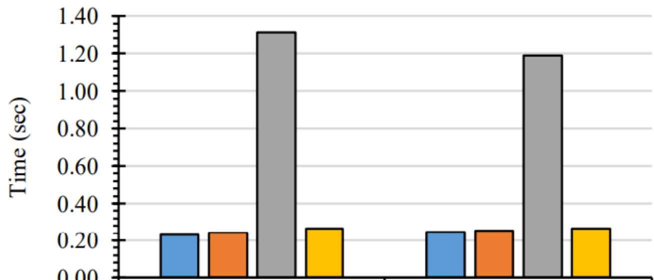
	Deeplabv3	Deeplabv3+
min	0.6139	0.5155
median	0.9634	0.9565
max	0.9879	0.9882
mean	0.9352	0.9202

Fig. 8. Overall Mean Intersection over Union



	Deeplabv3	Deeplabv3+
min	0.7695	0.7132
median	0.9835	0.9808
max	0.9948	0.9955
mean	0.9707	0.9632

Fig. 9. Pixel Accuracy



	Deeplabv3	Deeplabv3+
min	0.23	0.25
median	0.24	0.25
max	1.31	1.19
mean	0.26	0.27

Fig. 10. Inference Time per Image

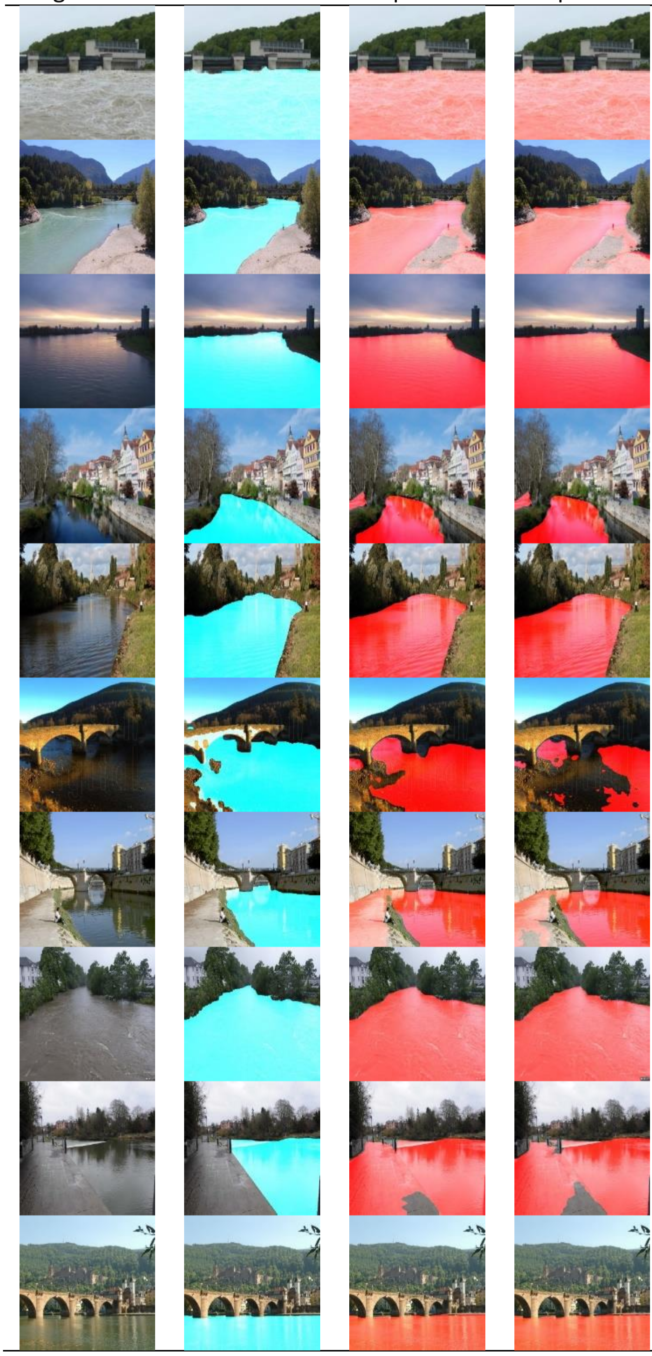
In terms of pixel accuracy, DeepLabv3 Mean value also indicates that the algorithm is better than DeepLabv3+ for segmentation of water river image with an accuracy of 97.07% as to Deeplabv3+ of 96.32%. Finally, the inference time per image Mean value also shown that the Deeplabv3 record a timing of 0.26 seconds which is 0.01 seconds faster than timing of 0.27 seconds for the Deeplabv3+. However, in terms of maximum value of accuracy obtained for Background MIoU, Overall MIoU, pixel accuracy and timing of inference time per image shown that DeepLabv3+ is better than DeepLabv3 with value of 99.41%, 98.82%, 99.55% and 1.19 seconds respectively.

3.4 Qualitative Visual Results Evaluation

Further look can be seen on qualitative result in Table 3 where DeepLabv3 algorithm succeed in classifying segmented part of water as DeepLabv3 seems to resemble more to ground truth. Based on the visualization results, water regions for ground truth are marked in blue colour while for DeepLabv3 and DeepLabv3+ they are marked in red colour in the images. In comparing both proposed method DeepLabv3 and DeepLabv3+, DeepLabv3 model are more likely able to further remove the isolated false positives and at the same time good at improving the prediction of the water region. Note that, the visualized results might seem to be similar for both DeepLabv3 and DeepLabv3+ in certain images, however for images with complex details and pixels, DeepLabv3 are seen to be more excellent in predicting the water image.

Table 3
Qualitative results of river segmentation, the first column refers to the original image, the second is the ground truth, the third is the result from DeepLabv3+ and the fourth one is the results from DeepLabv3

Original	GT	Deeplabv3+	Deeplabv3



4. Conclusions

In this paper, river segmentation with DeepLabv3 and DeepLabv3+ have been studied. This research work is conducted as the preliminary study in designing a better solution for monitoring the water level in river areas for an Early Flood Management System. Both algorithms are trained and implemented for river water image segmentation with Atrous Convolution. Results from this research shows that our developed method Deeplabv3 is the best among the other algorithms thus suit for the tasks of river water segmentation in both quantitatively and qualitatively. Thus, the implementation of our method is practical for monitoring the water level at riverbank using an image processing technique. For future work, the developed DeepLabv3 can be implemented on an edge computing device for real world applications and scenarios.

Acknowledgement
This research is funded by FRGS research grant – “Formulation of a New Density-Adaptive Classifier Framework based on Radiomics & U-Net Convolutional Neural Network (CNN) fused Features in Breast Lesion Mammogram” (File No: FRGS/1/2022/SKK06/UITM/02/3) from Ministry of Higher Education (MOHE), Malaysia. The author would like to express sincere gratitude to Universiti Teknologi MARA, Cawangan Pulau Pinang for permitting us to use their facility to complete this research work. Special thanks for those who helped directly and indirectly.

References
[1] Shaari, Nor Ashikin Binti. "Flood Causes Due to Some Geological Aspect: A case study in Kota Bharu, Kelantan, Malaysia." PhD diss., Faculty of Earth Science, 2016.
[2] Long, Jonathan, Evan Shelhamer, and Trevor Darrell. "Fully convolutional networks for semantic segmentation." In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 3431-3440. 2015. https://doi.org/10.1109/CVPR.2015.7298965
[3] Ciresan, Dan, Alessandro Giusti, Luca Gambardella, and Jürgen Schmidhuber. "Deep neural networks segment neuronal membranes in electron microscopy images." *Advances in neural information processing systems* 25 (2012).
[4] Ning, Feng, Damien Delhomme, Yann LeCun, Fabio Piano, Léon Bottou, and Paolo Emilio Barbano. "Toward automatic phenotyping of developing embryos from videos." *IEEE Transactions on Image Processing* 14, no. 9 (2005): 1360-1371. https://doi.org/10.1109/TIP.2005.852470
[5] Kang, Kai, and Xiaogang Wang. "Fully convolutional neural networks for crowd segmentation." *arXiv preprint arXiv:1411.4464* (2014).
[6] Marsden, Mark, Kevin McGuinness, Suzanne Little, and Noel E. O'Connor. "Fully convolutional crowd counting on highly congested scenes." *arXiv preprint arXiv:1612.00220* (2016). https://doi.org/10.5220/0006097300270033
[7] Lopez-Fuentes, Laura, Claudio Rossi, and Harald Skinnemoen. "River segmentation for flood monitoring." In *2017 IEEE international conference on big data (Big Data)*, pp. 3746-3749. IEEE, 2017. https://doi.org/10.1109/BigData.2017.8258373
[8] Isola, Phillip, Jun-Yan Zhu, Tinghui Zhou, and Alexei A. Efros. "Image-to-image translation with conditional adversarial networks." In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1125-1134. 2017. https://doi.org/10.1109/CVPR.2017.632
[9] Reed, Scott, Zeynep Akata, Xinchen Yan, Lajanugen Logeswaran, Bernt Schiele, and Honglak Lee. "Generative adversarial text to image synthesis." In *International conference on machine learning*, pp. 1060-1069. PMLR, 2016.
[10] Wu, Xian, Kun Xu, and Peter Hall. "A survey of image synthesis and editing with generative adversarial networks." *Tsinghua Science and Technology* 22, no. 6 (2017): 660-674. https://doi.org/10.23919/TST.2017.8195348
[11] Lebedev, M. A., Yu V. Vizilter, O. V. Vygolov, Vladimir A. Knyaz, and A. Yu Rubis. "Change detection in remote sensing images using conditional adversarial networks." *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 42 (2018): 565-571. https://doi.org/10.5194/isprs-archives-XLII-2-565-2018
[12] Huang, Gao, Zhuang Liu, Laurens Van Der Maaten, and Kilian Q. Weinberger. "Densely connected convolutional networks." In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4700-4708. 2017. https://doi.org/10.1109/CVPR.2017.243
[13] Jégou, Simon, Michal Drozdzal, David Vazquez, Adriana Romero, and Yoshua Bengio. "The one hundred layers tiramisu: Fully convolutional densenets for semantic segmentation." In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pp. 11-19. 2017. https://doi.org/10.1109/CVPRW.2017.156
[14] Hai, Jinjin, Kai Qiao, Jian Chen, Hongna Tan, Jingbo Xu, Lei Zeng, Dapeng Shi, and Bin Yan. "Fully convolutional densenet with multiscale context for automated breast tumor segmentation." *Journal of healthcare engineering* 2019 (2019). https://doi.org/10.1155/2019/8415485

[15] Jégou, Simon, Michal Drozdzal, David Vazquez, Adriana Romero, and Yoshua Bengio. "The one hundred layers tiramisu: Fully convolutional densenets for semantic segmentation." In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pp. 11-19. 2017. https://doi.org/10.1109/CVPRW.2017.156
[16] Lopez-Fuentes, Laura, Claudio Rossi, and Harald Skinnemoen. "River segmentation for flood monitoring." In *2017 IEEE international conference on big data (Big Data)*, pp. 3746-3749. IEEE, 2017. https://doi.org/10.1109/BigData.2017.8258373
[17] Chen, Liang-Chieh, George Papandreou, Florian Schroff, and Hartwig Adam. "Rethinking atrous convolution for semantic image segmentation." *arXiv preprint arXiv:1706.05587* (2017). https://doi.org/10.1007/978-3-030-01234-2_49
[18] Sunandini, Gosula, Ramesh Sivanpillai, V. Sowmya, and VV Sajith Variyar. "Significance of Atrous Spatial Pyramid Pooling (ASPP) in Deeplabv3+ for Water Body Segmentation." In *2023 10th International Conference on Signal Processing and Integrated Networks (SPIN)*, pp. 744-749. IEEE, 2023. https://doi.org/10.1109/SPIN57001.2023.10116882
[19] Li, Zhiqiang, Xi Chen, Jie Jiang, Zhen Han, Zhihong Li, Tao Fang, Hong Huo, Qingli Li, and Min Liu. "Cascaded multiscale structure with self-smoothing atrous convolution for semantic segmentation." *IEEE Transactions on Geoscience and Remote Sensing* 60 (2021): 1-13. https://doi.org/10.1109/TGRS.2021.3088902
[20] Wu, Jin, Yong Liu, Yuanpei Zhu, and Zun Li. "Atrous residual convolutional neural network based on U-Net for retinal vessel segmentation." *Plos one* 17, no. 8 (2022): e0273318. https://doi.org/10.1371/journal.pone.0273318