Flood Area Segmentation using Semantic Segmentation-based Deep Learning Models

Asghar Ali Chandio [1], Mehwish Leghari [2*], Harris [1], and Sahil Umar [1]

[1] Artificial Intelligence Department, Quaid-e-Awam University of Engineering Science & Technology, Nawabshah
[2] Data Science Department, Quaid-e-Awam University of Engineering Science & Technology, Nawabshah
* Correspondence: legharimehwish@quest.edu.pk

Abstract

Floods are natural disasters that frequently occur in different areas of Pakistan and cause several damages including human lives, infrastructure, and material losses. The unavailability of bench mark dataset of the flood data is the primary constraint in improving the flood monitoring systems. With state-of-the-art computer vision and deep learning techniques, flood detection, segmentation, and recognition systems have gained much attention. In this paper, a deep learning based DeepLabV3 architecture has been applied to segment the flood areas from the images. A ResNet-50 model has been used as a backbone network for extracting the features. The model has been trained and evaluated on the flood area segmentation dataset. The performance of the flood area recognition model has been assessed using the mean F1-score and Jaccard index. This research study also compares the performance of DeepLabV3 architecture with other backbone models including VGG-16 and DenseNet. Furthermore, the performance of UNet architecture has been evaluated and compared with DeepLabV3. Based on the experimental results, the DeepLabV3 architecture with ResNet-50 as a backbone model achieved the best segmentation results than the other models.

Keywords: Flood Area Segmentation; Semantic Segmentation; DeepLab Architecture; Flood Area Recognition

1. Introduction

The rapid change in the climate increases the vulnerability in flood across the globe. The climate change also continuously surges the frequency of the floods [1, 2]. Human factors also play a significant role in causing flood disasters, particularly through irrational deforestation and land use practices. The escalation of urbanization proliferates impermeable surfaces, which is disrupting natural drainage patterns and is a likelihood of flooding in urban areas [3]. Several countries, including Pakistan experience extreme flooding caused by heavy rain falls or seasonal downpours, which severely damage the infrastructure, economy and human lives [4–6]. Among various types of floods, including urban flooding, sewer flooding, flooding caused by glaciers and the coastal flooding, the river flooding often causes damages in the nearest areas as well as adjacent areas to rivers and has immense impact on the economy of the country [7]. Traditionally, a typical method for detecting floods involves using images from passive sensors like optical cameras. Various methods including the Normalized Difference Water Index (NDWI) and other water indices are often commonly applied. Although these techniques have shown promising results of flood detection, however, the imaginary data have lacking of cloud cover and the presence of extreme daylight, which significantly reduces the flood recognition or segmentation accuracy [7]. Recently, artificial intelligence particularly machine learning and deep learning techniques have shown promising results for flood and water detection, recognition, and segmentation [8–10]. In this paper four different deep learning-based models including VGG16 [11], DenseNet[12], DeepLabV3+[13], and UNet [14] are used to segment the flooded area from the images. Two different datasets of flood im- ages have been used to train and evaluate the performance of each deep-learning model. The datasets are publicly available at Kaggle [15, 16]. The datasets contain flood-hit areas along with the corresponding segmentation masks.

A deep learning-based pre-trained model is also used for VGG16, DenseNet, and DeepLabV3+ with ResNet50 as a backbone network. The model was trained on ImageNet dataset. The output of the last convolutional layer was taken and used as the input to the encoder module. For DeepLabV3+ model, the outputs of conv4 in block6 and conv2 in block3 were taken. These outputs contained high- level and low-level features, which were then concatenated. An upsampling layer for VGG16, DenseNet and DeepLabV3+ was used to make the dimensions of feature vectors equal. All the models were trained up to 100 epochs. Furthermore, an early stopping feature of Keras library was used to stop the training if the validation accuracy did not improve up to 5 epochs. The results obtained show that the DeepLabV3+ with ResNet50 as a backbone network segmented the flood areas from images more accurately than other models.

2. Methodology

The methodology used for different deep learning models to segment flood area from images is described in this section. It includes the collection of relevant datasets, data pre-processing steps, model training, and performance evaluation. Various deep learning architectures including VGG16, DenseNet, DeepLabV3+ and UNet were utilized and compared to achieve optimal results. The general methodology, as illustrated in Figure 1, outlines the common steps taken from data collection to model evaluation. Figure 2 illustrates the architecture of DeepLabV3+.

2.1. Data Collection

The datasets utilized for flood area segmentation were collected from Kaggle named Flood Area Segmentation, and flood dataset for semantic segmentation.

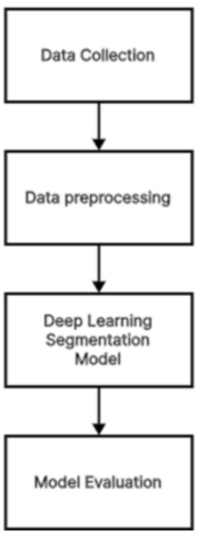

Figure 1. General Methodology used for flood area segmentation

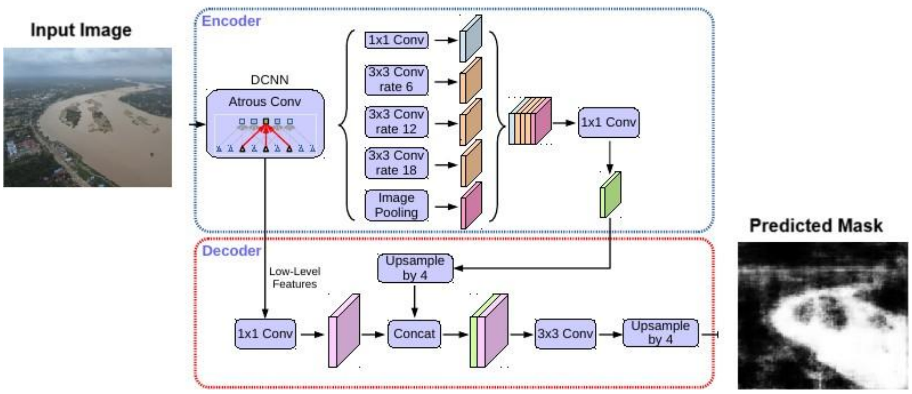

Figure 2. Architecture of DeepLabV3 with ResNet50 backbone network used for flood area segmentation from images[13]

These datasets comprised images of flood scenes along with their corresponding segmentation masks. The images represented varied flood conditions and environments, ensuring the dataset's applicability to real-world scenarios. We acknowledge the limitation of using publicly available datasets, which may not capture region-specific flood patterns such as those prevalent in Pakistan. To address this, future work will focus on collecting and annotating localized flood imagery datasets from affected areas in Pakistan to enhance model relevance and improve generalizability to regional flood conditions. The images used in this study are publicly available and anonymized. However, we acknowledge the importance of ethical considerations, particularly in disaster contexts. Future data collection will ensure compliance with privacy regulations, consent protocols, and ethical standards for capturing and using images of affected individuals and properties.

2.2. Data Preprocessing

To ensure the compatibility of the datasets with the segmentation models and enhance the training process, these preprocessing steps were carried out. **Image resizing:** All images and their corresponding masks were resized to a uniform dimension of (224x224x3) to standardize the input size across all the deep learning models. **Image Normalization**: The pixel data of the flood images are rescaled to the range of 0 and 1 by dividing each pixel value with 255. This step facilitated faster convergence during model training and ensured numerical stability. Furthermore, this method ensured the range of image pixel values between 0 and 1.

2.3. Deep Learning Segmentation Model Training

For training the segmentation model, the DeepLabV3+ architecture was utilized with ResNet-50 pre-trained on the ImageNet dataset as its backbone for feature extraction. The fully connected layers of the backbone ResNet50 were not included. The low-level and high-level features of convolutional layers in block3 and block6 were taken and concatenated. An upsampling2D was used to make the dimensions of the feature maps equal. For training an Adadelta optimizer with an initial learning rate of 0.001 was used. A binary cross-entropy loss function was applied to calculate the error value between the actual flood area masks and the predicted masks. Different batch sizes were used during the experiments; however, the best results were obtained with a batch size of 8. All the models were trained up to 100 epochs. The model was trained on flood area segmentation datasets to accurately segment flood-affected regions from input images. Preprocessing techniques were employed to enhance the quality of datasets, and the model training was performed using appropriate hyperparameters to achieve optimal results. Additionally, a comparative analysis was conducted by training the DeepLabV3+ architecture with other backbone models, including VGG- 16 and DenseNet, and by evaluating the performance of the UNet architecture. This study was conducted using a system with a NVIDIA T4 GPU, 16GB RAM, and an Intel Core i7 processor. Training time varied by model complexity, with DeepLabV3+ requiring approximately 4 hours for 100 epochs.

2.4. Model Evaluation

The evaluation of the trained segmentation models was conducted to assess their performance in accurately segmenting flood-affected areas. The performance of each model was measured using key metrics, including the Jaccard Similarity Score (also known as Intersection over Union), which quantified the overlap be- tween the predicted segmentation and the ground truth, reflecting the accuracy of the model's segmentation and the mean F1 score. The models were tested on a diverse set of flood images to determine how well they generalized to real-world scenarios. A comparative analysis was carried out to benchmark the DeepLabV3 architecture with ResNet50 against other models like VGG-16, DenseNet, and UNet.

3. Results and Discussion

Table 1 shows the mean F1 score and Jaccard similarity score of four different deep learning models. As shown in Table 1 the mean F1 score and the Jaccard similarity score of DeepLabV3+ with ResNet50 as a backbone network is better than DenseNet, VGG16 and UNet models. The mean F1 score and the Jac- card similarity score were calculated using the built-in functions implemented in Scikit-Learn machine learning library. The DeepLabV3+ model implemented Atrous Spatial Pyramid Pooling to extract valuable information from the activation maps of the ResNet50 model at various scales. Both the high-level and low- level features are extracted for better understanding of the flood areas present in the data, whereas to extract the location-based features from the flood images, it takes activations maps from the initial layers of the ResNet50. The model later concatenates both the high-level and low-level features for predicting more accurate flood masks.

Table 1. Deep Learning Model Performance Comparison

Model Name	Mean F1 Score	Jaccard Similarity Score
DenseNet	0.80	0.65
VGG16	0.74	0.54
UNet	0.78	0.63
DeepLabV3+	0.82	0.66

3.1. Training vs. Validation Accuracy and Losses

The training and validation accuracies as well as the training and validation loss values of all the deep learning models were measured using evaluation protocols. The best results of training and validation accuracies of the two models are illustrated in Figure 3 and Figure 4.

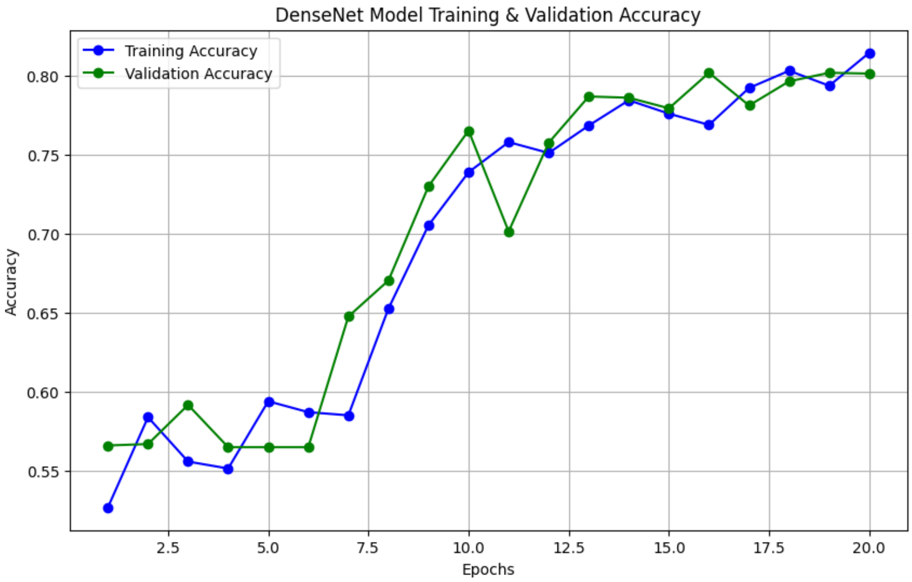

Figure 3. Training and Validation Accuracy Visualization of DenseNet

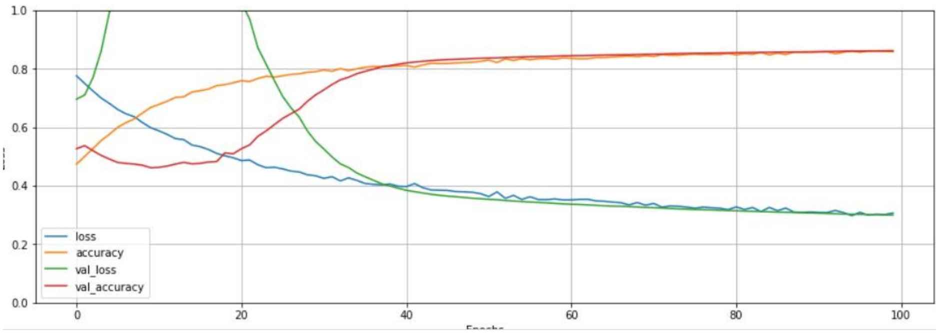

Figure 4. Training and Validation Accuracy and Loss Visualization of DeepLabV3+

Figure 5. Visualisation of DenseNet Results

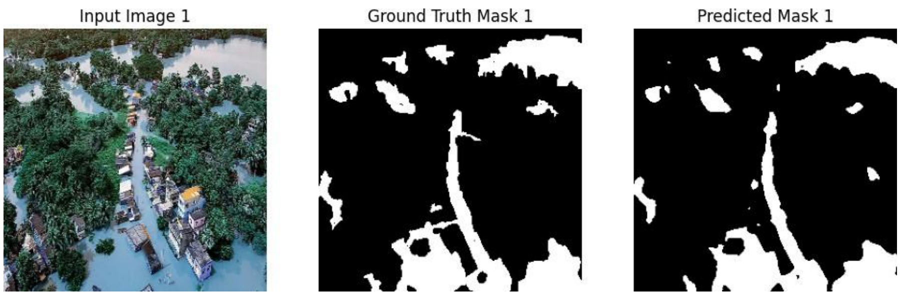

Figure 6. Visualisation of Unet Results

Figure 7. Visualization of DeepLabV3+ Results

4. Conclusion And Future Work

In this paper, four different models of deep learning were used to segment the flood areas from the image data. The models used include VGG16, DenseNet, UNet and the DeepLabeV3+. The DeepLabV3+ model was trained with different backbone networks, however, the best segmentation results were achieved with the ResNet50 network. The DeepLabV3+ model uses Atrous Spatial Pyramid Pooling layer to extract the useful features from the activation maps of the ResNet50 model at various scales. Both the high-level and low-level features are extracted and concatenated to accurately predict the segmentation masks. The performance of each deep learning model in terms of mean F1 score and Jaccard similarity score was evaluated on two different datasets of flood area images containing both the flood images and their segmented masks. The datasets used are publicly available at Kaggle. In the future, the real data of flood area images from Pakistani regions will be collected to train the state-of-the-art deep learning models and evaluate their segmentation predictions. Furthermore, the instance segmentation method will be employed to classify the flood image pixels into flood and non-flood regions more accurately. While this study evaluates traditional and popular CNN-based segmentation models (DeepLabV3+, UNet, DenseNet, VGG16), we recognize the growing potential of transformer-based models such as Vision Transformers (ViT) and Swin Transformers. As part of future research, we plan to explore and benchmark these state-of-the-art architectures for flood segmentation tasks to determine performance improvements over CNN-based approaches.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Kim, J., Kim, H., Kim, D. J., Song, J., & Li, C.: Deep Learning-Based Flood Area Extraction for Fully Automated and Persistent Flood Monitoring Using Cloud Computing. Remote Sensing, 14(24), 6373 (2022).
2. Pally, R. J., & Samadi, S.: Application of image processing and convolutional neural networks for flood image classification and semantic segmentation. Environmental modelling & software, 148, 105285 (2022).
3. Wang, Y., Shen, Y., Salahshour, B., Cetin, M., Iftekharuddin, K., Tahvildari, N., & Goodall, J. L.: Urban flood extent segmentation and evaluation from real-world surveillance camera images using deep convolutional neural network. Environmental Modelling & Software, 173, 105939 (2024).
4. Priscillia, S., Schillaci, C., & Lipani, A.: Flood susceptibility assessment using artificial neural networks in Indonesia. Artificial Intelligence in Geosciences, 2, 215-222 (2021).
5. Essam, Y., Huang, Y. F., Ng, J. L., Birima, A. H., Ahmed, A. N., & El-Shafie, A.: Predicting streamflow in Peninsular Malaysia using support vector machine and deep learning algorithms. Scientific Reports, 12(1), 3883 (2022).
6. Lemenkova, P.: Deep Learning Methods of Satellite Image Processing for Monitoring of Flood Dynamics in the Ganges Delta, Bangladesh. Water, 16(8), 1141 (2024).

7. Nemni, E., Bullock, J., Belabbes, S., & Bromley, L.: Fully convolutional neural network for rapid flood segmentation in synthetic aperture radar imagery. Remote Sensing, 12(16), 2532 (2020).
8. Fu, G., Jin, Y., Sun, S., Yuan, Z., & Butler, D.: The role of deep learning in urban water management: A critical review. Water Research, 223, 118973 (2022).
9. Munawar, H. S., Ullah, F., Qayyum, S., & Heravi, A.: Application of deep learning on uav-based aerial images for flood detection. Smart Cities, 4(3), 1220-1242 (2021).
10. Bahrami, B., & Arbabkhah, H.: Enhanced Flood Detection Through Precise Water Segmentation Using Advanced Deep Learning Models. Journal of Civil Engineer- ing Researchers, 6(1), 1-8 (2024).
11. Simonyan, K.: Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556, (2014).
12. Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q.: Densely connected convolutional networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4700-4708) (2017).
13. Chen, L. C.: Rethinking atrous convolution for semantic image segmentation. arXiv preprint arXiv:1706.05587 (2017).
14. Ronneberger, O., Fischer, P., & Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18 (pp. 234-241). Springer International Publishing (2015).
15. Kaggle Homepage, https://www.kaggle.com/datasets/faizalkarim/flood-area- segmentation, last accessed 2024/12/05
16. Kaggle Homepage, https://www.kaggle.com/datasets/warcoder/flood-dataset-for- semantic-segmentation, last accessed 2024/12/05