# Analysing Flood Impact: Semantic Segmentation of Imagery with FloodNet Dataset

**UEC2021155 Siddhi Patil**
**UEC2021162 Shreeparna Dhara**
**UEC2021166 Charul Wankhede**

## ABSTRACT

Semantic segmentation of remote sensing images is a vital task with numerous applications, including environmental monitoring, disaster management, and urban planning. In this paper, we propose a deep learning approach for semantic segmentation using the Floodnet dataset, which comprises large-scale remote sensing images in the TIFF format. Due to the dataset's size and the computational demands, we adopted a strategy of converting the images into NumPy arrays for efficient handling. We utilized a UNet architecture based on the VGG16 convolutional neural network (CNN) as the backbone for semantic segmentation. The VGG16 network, pre-trained on ImageNet, provides a robust feature extractor, which is crucial for capturing high-level features in remote sensing images. We describe the data preprocessing steps, including data augmentation techniques to enhance the model's generalization ability. Additionally, we present the implementation details, including model architecture, training procedure, and evaluation metrics. The performance of the proposed approach is evaluated on a test set, and quantitative metrics such as accuracy and Intersection over Union (IoU) are reported. We demonstrate the effectiveness of the proposed method through qualitative analysis by visualizing segmentation results on sample images from the test set.

## I. INTRODUCTION

Semantic segmentation of remote sensing images plays a pivotal role in extracting valuable information for various applications, including environmental monitoring, disaster management, and urban planning. In this study, we propose a novel approach to semantic segmentation using the FloodNet dataset, a curated collection of remote sensing images specifically tailored for flood detection and monitoring. Leveraging the capabilities of deep learning and the U-Net architecture with features from the VGG16 network, our method aims to accurately delineate semantic regions within flood-affected areas. By combining the rich spatial and spectral information present in remote sensing imagery with advanced deep learning techniques, we seek to address the challenges associated with flood detection and mapping in diverse environmental conditions. This research contributes to the advancement of semantic segmentation methodologies tailored for remote sensing applications, ultimately facilitating more effective flood monitoring and mitigation strategies.

## II. METHODOLOGY

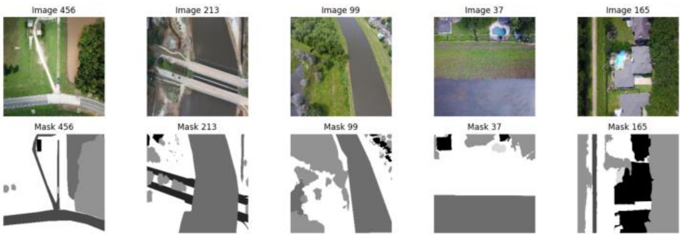
### 2.1 Image Processing

We initialize our research by accessing and loading essential data resources crucial for our investigation into semantic segmentation of remote sensing images. Leveraging the powerful NumPy library, we load two fundamental arrays from files stored in a designated directory within Google Drive. These arrays, denoted as train_images and train_masks, respectively, contain preprocessed images and their corresponding masks.

The images are derived from the Floodnet dataset, a substantial repository of remote sensing imagery.

## 2.2 Floodnet Dataset

The FloodNet dataset is a curated collection of remote sensing images specifically tailored for flood detection and monitoring applications. It comprises diverse imagery types, including satellite, aerial, and UAV images, captured during various flood events, offering comprehensive coverage of flood-affected areas. Each image in the dataset is accompanied by corresponding annotations, delineating the extent of flooded regions, enabling supervised learning approaches for flood detection and segmentation tasks. With high spatial resolution and temporal coverage spanning multiple flood events, the FloodNet dataset facilitates detailed analysis of flood dynamics and impacts.




## 2.3 Data Augmentation

We randomly select images from the training dataset and apply transformations such as rotation, flipping, and scaling to generate augmented samples, effectively increasing the diversity and variability of the training data. Visualization of augmented samples provides insights into the transformation effects on both input images and corresponding masks, aiding in the understanding of the augmentation process. Furthermore, we expand the label space by converting the segmentation masks into categorical format using one-hot encoding, facilitating compatibility with deep learning frameworks. The augmented dataset, comprising original and transformed samples, is then split into training and validation sets for model training and evaluation.

## 2.4 Model

Encoder (Feature Extraction):
The encoder component of the model capitalizes on the convolutional layers from the pre-trained VGG16 network, deliberately excluding its fully connected layers. By leveraging the convolutional layers, particularly focusing on the 'block5_conv3' layer of the VGG16 architecture, the model adeptly captures high-level semantic features inherent in the input images. This process enables the extraction of abstract representations from the images, facilitating the subsequent segmentation tasks by preserving essential semantic information.

Decoder (Feature Expansion):
In contrast, the decoder section of the model is tasked with expanding the extracted features to reconstruct the original spatial dimensions of the input images. Comprising a sequence of convolutional and transposed convolutional layers, the decoder progressively upsamples the feature maps obtained from the encoder. Furthermore, it employs skip connections, which concatenate feature maps from corresponding layers in the encoder. This strategic integration of skip connections serves to preserve spatial information and facilitate precise localization of semantic regions within the images, enhancing the segmentation accuracy.

Output Layer:
At the culmination of the model architecture lies the output layer, a pivotal component responsible for producing pixel-wise class probabilities. Comprising a single 1x1 convolutional layer with softmax activation, this layer generates class probabilities for each pixel in the image, corresponding to the specified number of classes. The softmax activation function ensures that the output probabilities sum up to one for each pixel, thereby facilitating the interpretation of the model's predictions as class probabilities, crucial for accurate semantic segmentation.

Model Creation and Compilation:
The create_unet_vgg16 function orchestrates the assembly of the U-Net model architecture, seamlessly connecting the encoder and decoder components. Upon construction, the model undergoes compilation, a vital step facilitated by the Adam optimizer and categorical cross-entropy loss function. Tailored for multi-class segmentation tasks, this choice of optimizer and loss function ensures efficient model training and convergence. Additionally, accuracy is designated as the evaluation metric to gauge the model's performance throughout the training and validation phases, providing insights into its ability to accurately delineate semantic regions in remote sensing imagery

## III. **RESULTS**

### A. **Figures and Tables**

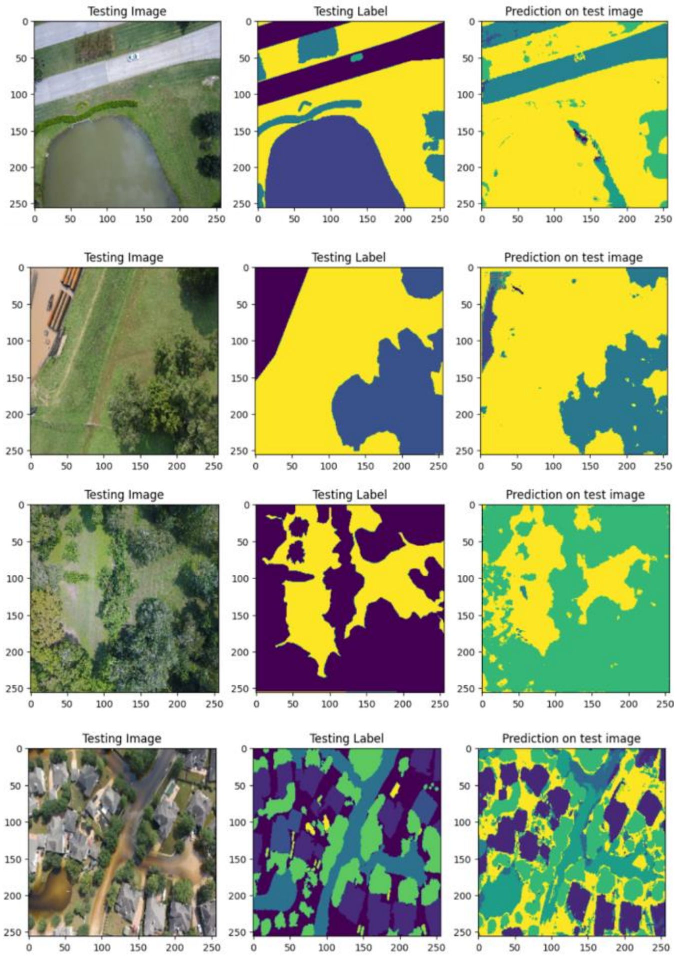
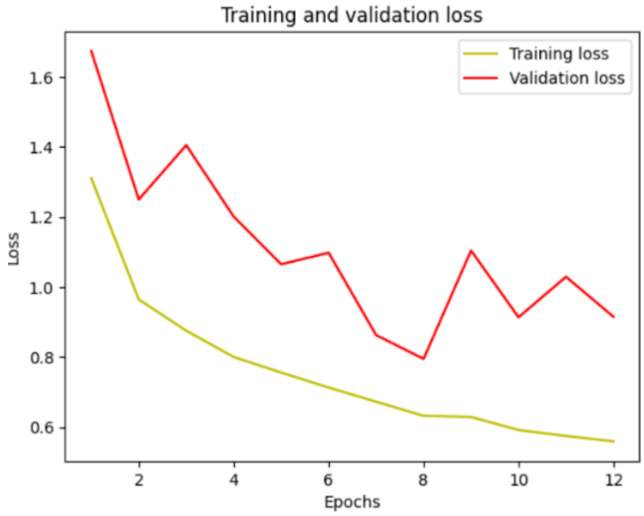



Figure 1: Actual images, Ground truths, Predicted

Loss = Categorical Entropy

| Optimizer | Learning rate | Data Augmentation | Accuracy | Mean IOU |
|---|---|---|---|---|
| Adam | 0.001 | Yes | 58.41 | 0.17 |
| Adam | 0.002 | No | 63.33 | 0.211 |
| SGD | 0.001 | Yes | 59.08 | 0.17 |
| SGD | 0.001 | Yes | 64.80 | 0.25 |
| **RMS Prop** | **0.002** | **No** | **73.33** | **0.47** |
| RMS Prop | 0.001 | Yes | 68.08 | 0.29 |

The resemblance between the predicted segmentation and the ground truth label on unseen data showcases the model's capacity to effectively capture intricate patterns and semantics within images. This outcome not only validates the efficacy of our model architecture but also highlights its potential for real-world applications where accurate pixel-level classification is imperative.




## IV. CONCLUSION

. In summary, our study presents a novel approach for semantic segmentation of flood-affected areas using the FloodNet dataset. Leveraging deep learning techniques and the U-Net architecture with features from the VGG16 network, we achieved accurate delineation of semantic regions in remote sensing imagery. Through meticulous data preprocessing and augmentation, our model demonstrated robust performance, validated by both quantitative metrics and qualitative visualization of segmentation results. Our methodology contributes to advancing flood monitoring and mitigation efforts, offering insights for environmental management and disaster response. Further refinements and extensions, including multi-temporal data fusion, hold promise for enhancing the applicability of our approach in real-world scenarios.

## V. REFERENCES

[1] M. Rahnemoonfar, T. Chowdhury, A. Sarkar, D. Varshney, M. Yari and R. R. Murphy, "FloodNet: A High Resolution Aerial Imagery Dataset for Post Flood Scene Understanding," in *IEEE Access*, vol. 9, pp. 89644-89654, 2021, doi: 10.1109/ACCESS.2021.3090981

[2] Chaoyun Mai, Yibo Wu, Yikui Zhai, Hao Quan, Jianhong Zhou, Angelo Genovese, Vincenzo Piuri, Fabio Scotti, "DBCG-Net: Dual Branch Calibration Guided Deep Network for UAV Images Semantic Segmentation", *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol.17, pp.7932-7945, 2024

[3] Abhisek Maiti, Sander Oude Elberink, George Vosselman, "UAVPal: A New Dataset for Semantic Segmentation in Complex Urban Landscape With Efficient Multiscale Segmentation", *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol.17, pp.464-475, 2024