Image Segmentation Models for Flood Inundiation Mapping

Aolong Li, Sean Dixit, Vishakha Dikshit, Yevhen Melnyk

aolli@iu.edu, sedixit@iu.edu, vidiksh@iu.edu, ymelnyk@iu.edu

Introduction

- Floods are one of the most devastating natural disasters, affecting over 20% of the world's population.
- In the framework of active learning for flood inundation mapping, we experiment different backbone models for flood pixel segmentation.
- We evaluate the performance of several GeoAI models (U-Net, Prithvi, Prithvi+Unet, Segformer) to evaluate their suitability and potential advantages for flood inundation mapping.

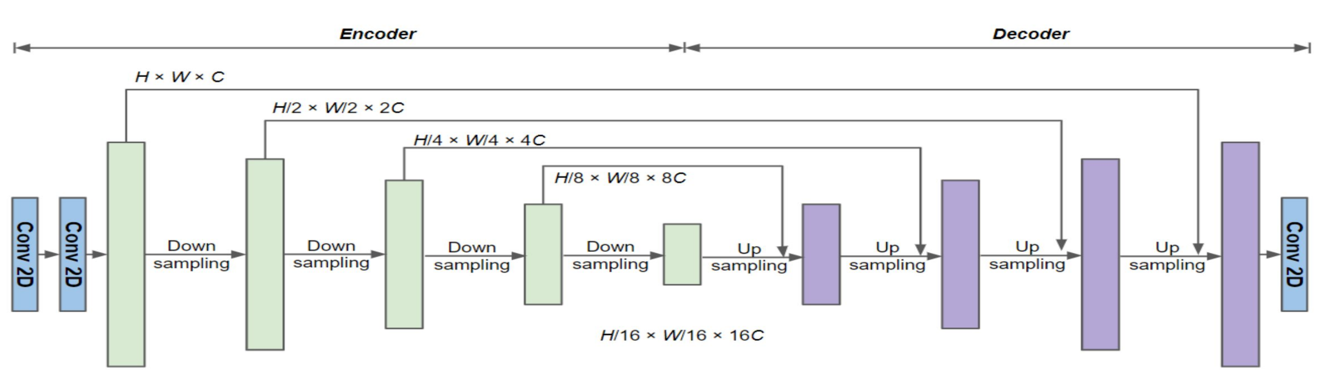

Fig 1. Structure of U-Net Model

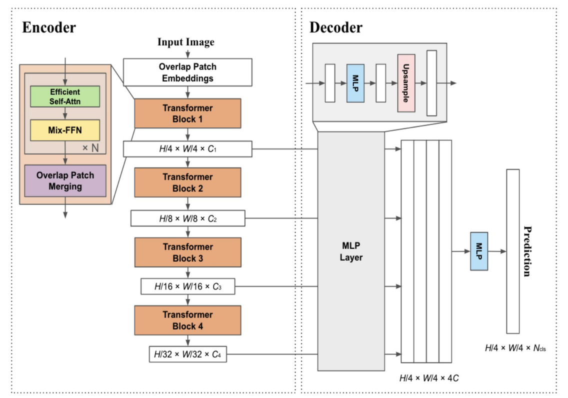

Fig 2. Structure of Segformer

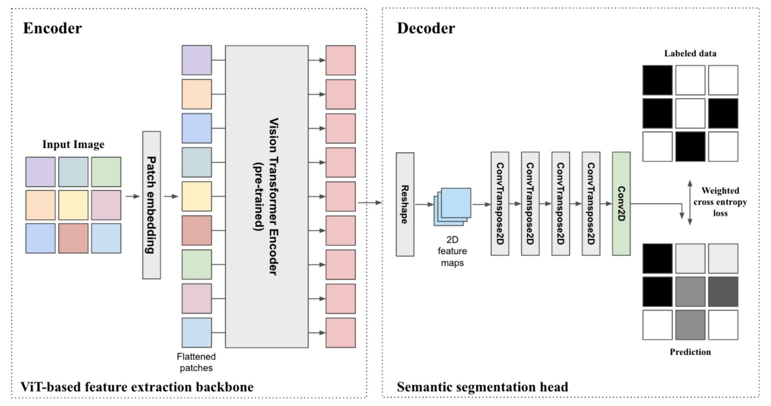

Fig 3. Structure of Prithvi

Methods

- **Data Preparation:**

  Our analysis uses satellite images of regions captured before and after flood events provide by Saugat Adhikari. Each pixel in these images is described by 7 channels, including RGB and elevation data, although the elevation data is not currently utilized in the model, Prithvi. To facilitate machine learning analysis, the images are padded to ensure they can be divided into fixed-size patches, allowing consistent and efficient processing.

- **U-Net:**

  Encoder (downsampling path): Convolutional blocks that progressively downsample the input, extracting increasingly abstract features. Each block applies convolutions, batch normalization, ReLU activation, and max pooling.

  Decoder (upsampling path): Convolutional blocks that progressively upsample the feature maps to the original image size, aiming to localize and delineate the object boundaries precisely. This ends with a convolution layer that maps the features to the desired number of classes.

  We experimented with several U-net variations, including a shallower encoder and decoder, and implemented upsampling within the decoder.

- **Segformer**

  Hierarchical Transformer Encoder: Uses the MiT (Mix Transformer) backbone which processes image features at various scales through multiple stages. Overlapping patch embedding and efficient self-attention mechanisms allow for detailed feature extraction across different resolutions.

  All-MLP Decoder: Simple MLP-based decoder that fuses multi-scale features from the encoder and predicts segmentation masks

- **Prithivi-Unet** (prithvi encoder + Unet decoder)

  Utilizes a transformer-based approach (specifically a Masked Autoencoder ViT) for robust and scalable feature extraction with Unet Decoder

- **Prithvi** (Encoder and Decoder)

  Along with Masked Autoencoder ViT as Encoder uses the Prithvi Encoder, transforming the rich encoded features back into spatial data for detailed image reconstruction or segmentation. Utilizes a combination of convolutional layers and transformer tokens to map deep features to output tasks, enhancing the accuracy of final predictions.

Results

Table 1: Results on unseen satellite imagery

Model	Avg. mIoU (%)	Avg. IoU (%)		Avg. mAcc (%)	Avg. Acc (%)		Number of trainable parameters
		Flood	Non-flood		Flood	Non-flood	
U-Net	**87.02**	**88.21**	**92.48**	**90.38**	**92.48**	**98.12**	29M
IBM-NASA's Prithvi	80.24	82.98	83.12	85.12	85.98	95.45	100M
Prithvi + U-Net	60.49	72.19	72.13	75.09	73.29	97.21	100M
Segformer	72.97	78.01	79.02	81.02	79.05	96.12	5M

Fig 4. Satellite image of flooded area

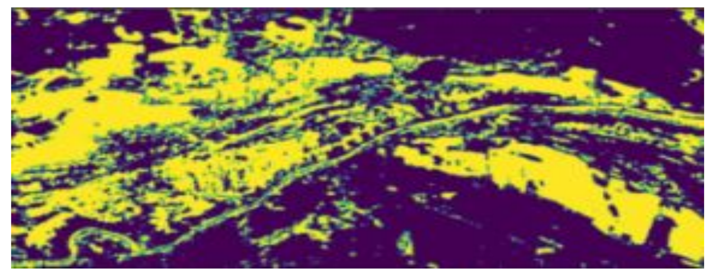

Fig 5. Segmented image using U-Net

Fig 6. Segmented image using IBM-NASA's Prithvi

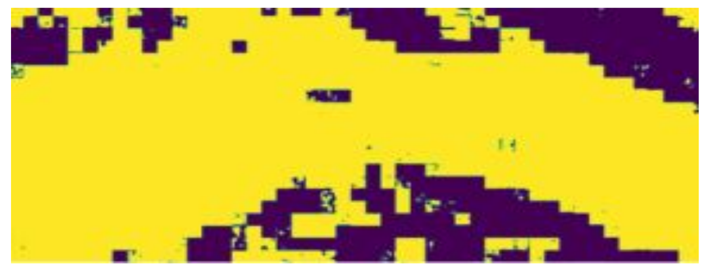

Fig 7. Segmented image using Prithvi + Unet

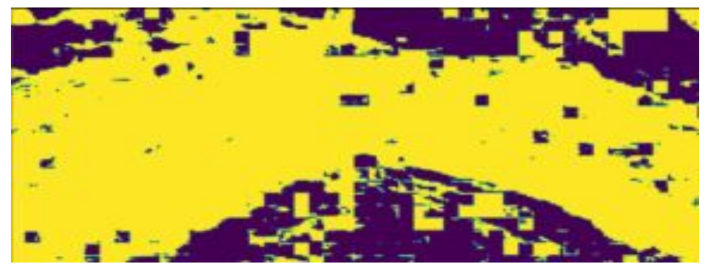

Fig 8. Segmented image using Segformer

Discussion and Future Work

- UNet outperformed all other models in flood inundation mapping, even with fewer parameters than the Prithvi model.
- Prithvi encoder with UNet decoder showed the poorest performance among the tested models.
- Further optimization of the existing models through hyperparameter tuning to potentially enhance performance.
- Evaluation of the DeepLab and Stacked Hourglass models to assess their efficacy in flood inundation mapping compared to the existing models.
- Obtaining more annotated satellite imagery data with elevation data to allow for more training and robustness of models.

References

- Li, Wenwen, et al. "Assessment of IBM and NASA's geospatial foundation model in flood inundation mapping." arXiv preprint arXiv:2309.14500 (2023).
- Isaaccorley. (n.d.). Isaaccorley/Prithvi-Pytorch. GitHub. https://github.com/isaaccorley/prithvi-pytorch

Acknowledgements

- We would like to thank professor Ya Dan and Saugat Adhikari for the project idea and their assistance throughout the project
- We would also like to express our gratitude to Professor David Crandall for his excellent teaching