Segment Any Stream (SAS): Scalable Water Extent Detection with the Segment Anything Model

Haozhen Zheng[1,3], Chenhui Zhang[1,3,4*], Kaiyu Guan[1,2,3], Yawen Deng[1,2], Sherrie Wang[4], Bruce L Rhoads[5], Andrew J Margenot[1,2], Shengnan Zhou[1,2], Sheng Wang[1,2*]

1. Agroecosystem Sustainability Center, Institute for Sustainability, Energy, and Environment, University of Illinois Urbana-Champaign, Urbana, IL 61801, USA
2. College of Agricultural, Consumer and Environmental Sciences, University of Illinois Urbana-Champaign, Urbana, IL 61801, USA
3. National Center for Supercomputing Applications, University of Illinois Urbana-Champaign, Urbana, IL 61801, USA
4. MIT Institute for Data, Systems, and Society, 77 Massachusetts Avenue, Cambridge, MA 02139, USA
5. Department of Geography & Geographic Information Science, University of Illinois Urbana-Champaign, Urbana, IL 61801, USA

***Corresponding: sheng12@illinois.edu, chenhui5@mit.edu**

Introduction

- Stream and river bank erosions are overlooked sources for soil erosions in the US Midwest and around the world (Zhou et al., 2022).
- The rapid progress in self-supervised learning foundation models, e.g., the Segment Anything Model (Kirillov et al., 2023), provides a potential tool to analyze high-resolution remote sensing imagery to efficiently detect water extent to quantify stream/river bank erosion rates.

Two scientific questions:

- What is the performance of the image segmentation foundation model, SAM, to identify river water extent from high-resolution remote sensing imagery?
- How do we properly fine-tune the foundation model SAM, which has a large number of parameters, for remote sensing tasks?

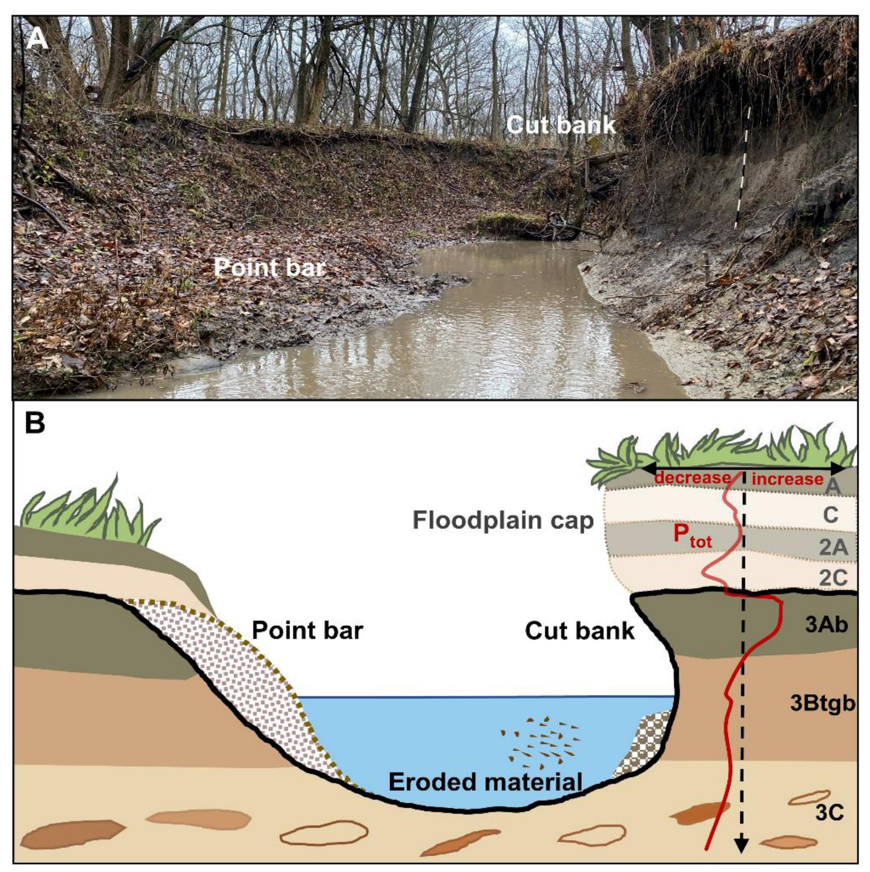

Fig 1. Conceptual diagram of stream/river bank erosion in a meander bend of Polecat Creek, IL, US.

Study region

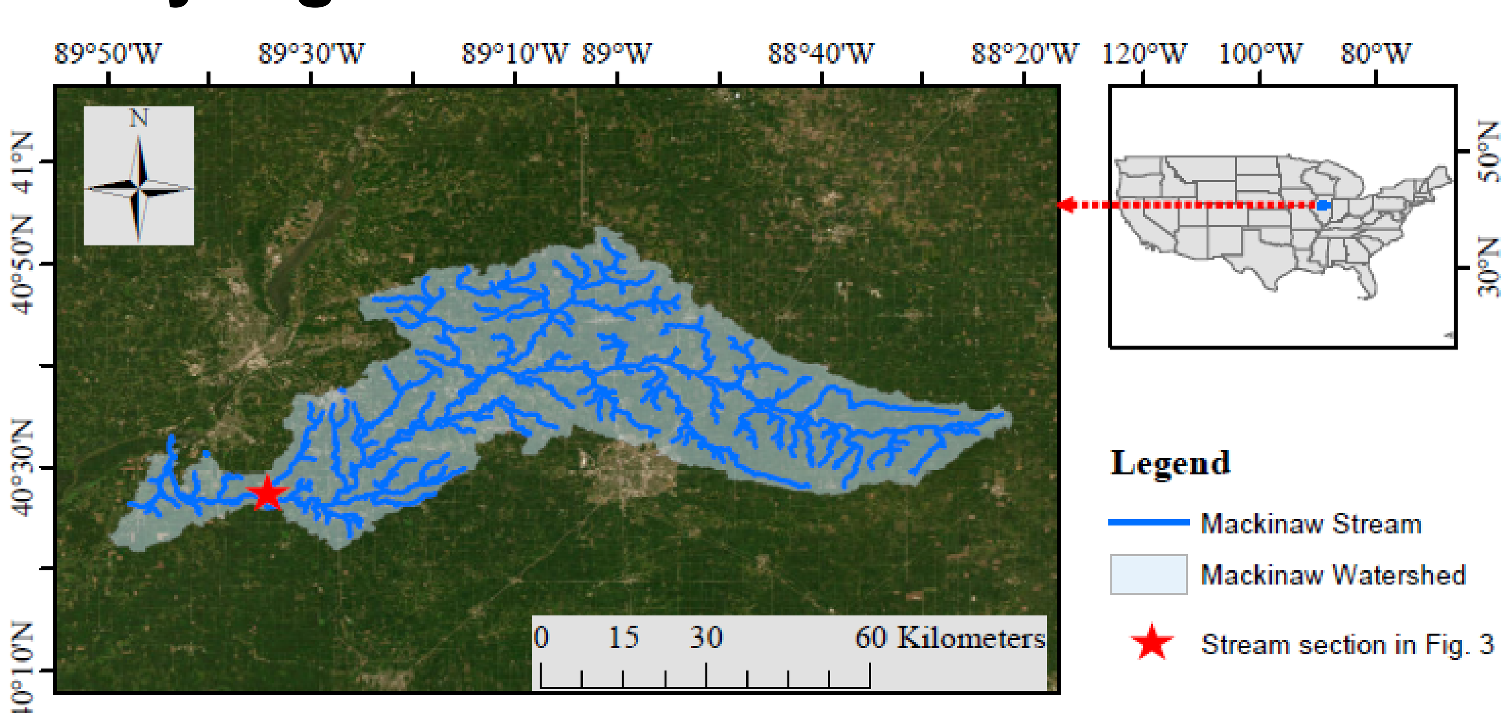

Fig 2. Study region of Mackinaw River watershed in central IL, US.

Methods and Data

We thoroughly compared and evaluated six approaches to detect water extent in streams and rivers.

1. Normalized Difference Water Index (NDWI) with kernel density estimation (KDE) for automatic thresholding
2. NDWI with Otsu for automatic thresholding
3. Supervised learning, U-Net
4. SAM-D: default SAM
5. SAM-P: point prompts from NDWI mask but untuned SAM
6. Segment Any Stream/river (SAS): SAM with LoRA for model fine-tuning

Data: USGS and USDA high-resolution airborne imagery (1m)
USGS data: leaf-off imagery collected in the non-growing season
USDA data: leaf-on imagery collected in the growing season

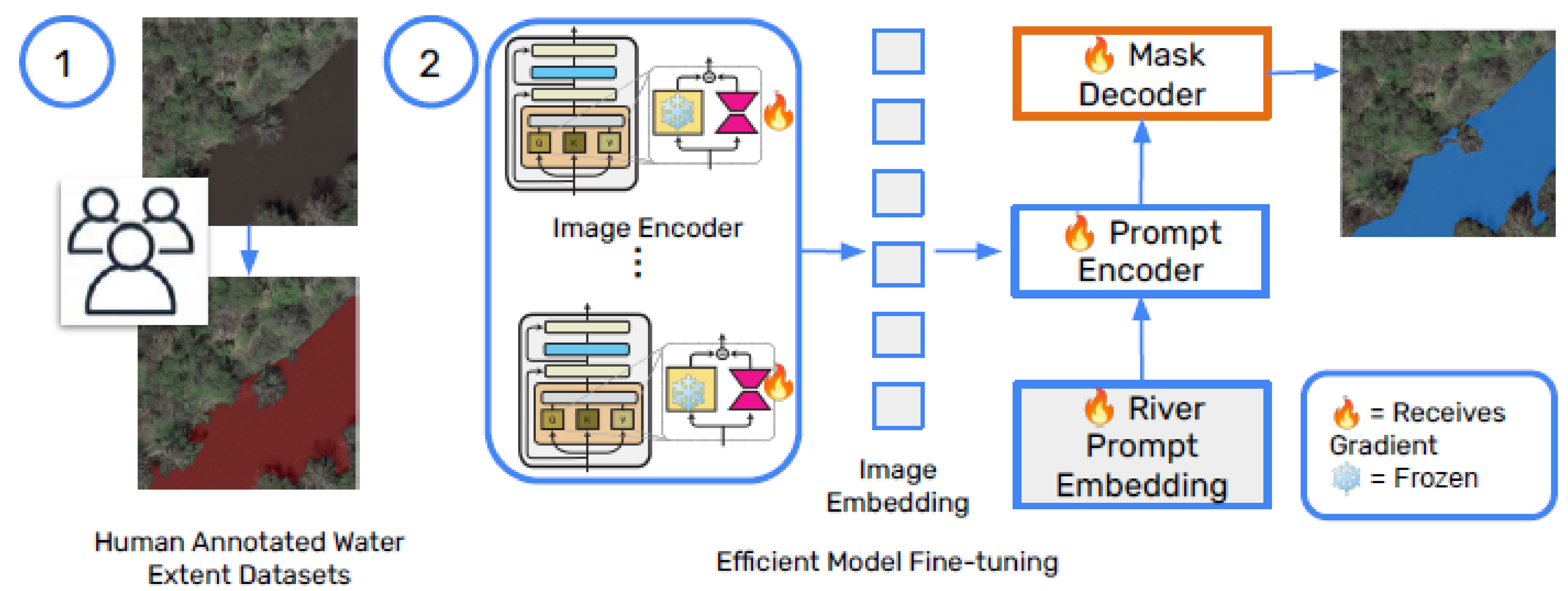

Fig. 3. Proposed the **Low-Rank Adaptation (LoRA)** method to fine-tune a pre-trained SAM in an efficient manner. LoRA constrains the weight updates to a low-rank subspace, facilitating a lightweight yet effective adaptation of SAM, circumventing the computational exigencies typically associated with full-scale fine-tuning.

Results

Table 1. Testing performance of different approaches in the study region compared to human labeling. These methods include spectral index thresholding, U-Net, SAM-D, SAM-P, and SAS. The best metrics were highlighted in bold.

Method	Input Bands	IoU	Precision	Recall	Accuracy	f-1	Kappa
NDWI (KDE)	G, NIR	0.35	0.55	0.42	0.86	0.78	0.68
NDWI (Otsu)	G, NIR	0.35	0.55	0.42	0.83	0.78	0.64
U-Net	RGB	0.71	0.87	0.79	0.89	0.83	0.75
SAM-D	RGB	0.17	0.33	0.26	0.57	0.29	-0.01
SAM-P	RGB	0.54	0.66	0.75	0.78	0.70	0.53
SAS (ours)	RGB	**0.76**	**0.90**	**0.83**	**0.90**	**0.86**	**0.79**

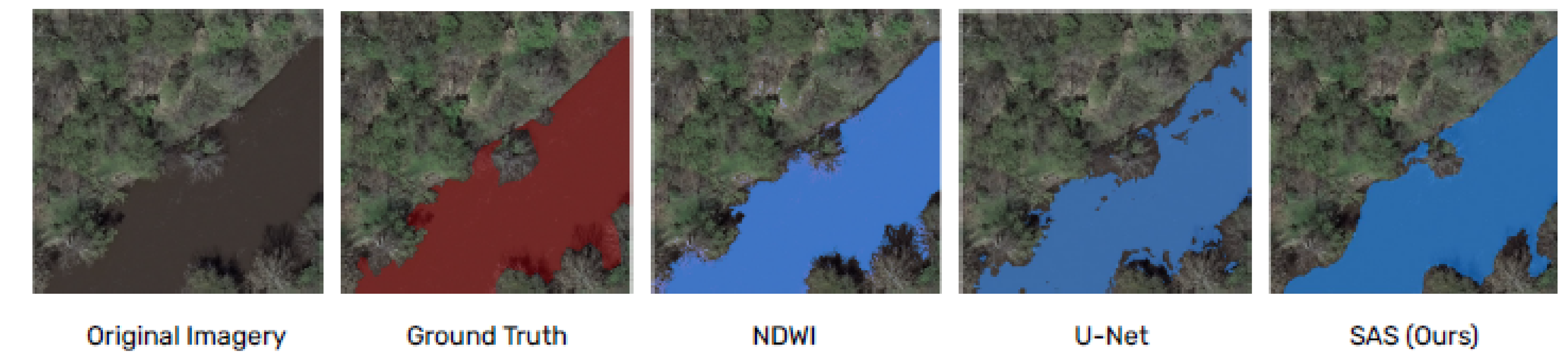

Fig. 4 Visualization of the comparison among different methods

Table 2. Testing performance of SAS in the study region with different LoRA rank on segmentation results. We performed an ablation test with LoRA ranks from 1 to 8. We did not observe a significant change in testing results with varying rank numbers.

Rank	% of Parameters Trained	IoU	Precision	Recall	Accuracy	f-1	Kappa
1	4.19	0.75	**0.89**	**0.82**	**0.91**	**0.84**	0.77
2	4.23	0.74	**0.89**	**0.82**	0.90	0.83	0.76
4	4.31	0.75	0.88	**0.82**	0.88	0.82	**0.78**
6	4.39	0.74	0.88	**0.82**	**0.91**	**0.84**	**0.78**
8	4.79	**0.75**	**0.89**	**0.82**	**0.91**	**0.84**	0.77

Conclusion

- In this work, we propose Segment Any Stream/River (SAS), a parameter-efficient fine-tuning framework that enables the adaptation of a state-of-the-art SAM model to segment river/stream water extent with higher accuracies than traditional spectral indices and supervised learning approaches.
- We also curate a regional watershed-scale water extent dataset in Illinois based on our annotations of the Mackinaw watershed. Through our experiments, we show that SAS demonstrates superior performance with low computational needs (consumer-grade GPU within two hours).

References

Zhou, S., Li, N. and Margenot, A.J., 2022. Soil meets stream: Vertical distribution of soil phosphorus in streambanks. Geoderma, 424, p.115989.

Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A.C., Lo, W.Y. and Dollár, P., 2023. Segment anything. arXiv preprint arXiv:2304.02643.

Acknowledgment

ILLINOIS NREC
NASA
DISCOVERY PARTNERS INSTITUTE
PART OF THE UNIVERSITY OF ILLINOIS SYSTEM