Beyond Surface Mapping: A Multimodal Pipeline for 3D Flood Depth Estimation with Foundation Models

Wenfeng Jia[1], Bin Liang[2], Muhammad Arif Khan[1], and Lihong Zheng[1]

[1] Charles Sturt University, New South Wales, Australia
[2] University of Technology Sydney, New South Wales, Australia

1 Introduction

Floods are one of the most frequent natural hazards, posing significant threats to human safety and critical infrastructure. Flood mapping technologies have emerged as essential tools for disaster mitigation, urban planning, and emergency response. While traditional 2D flood mapping effectively delineates inundation extents, it fails to capture crucial volumetric characteristics such as water depth - a critical parameter for floods.

Hydrodynamic models (e.g., TUFLOW, HEC-RAS, MIKE 21) based on shallow water equations have dominated 2D flood simulation due to their physical interpretability. However, extending these models to 3D applications introduces computational bottlenecks and scalability limitations. Therefore, some scholars have employed physics-constrained machine learning to conduct flood research[3]. With the advancement of deep learning (DL), other researchers have turned to DL to address these challenges. Zhou et al.[4] developed a spatial reduction and reconstruction framework based on LSTM, achieving two orders of magnitude speed improvements over the hydrodynamic model. Some approaches combine 2D mapping with digital elevation models (DEM) to achieve 3D flood mapping; Gebrehiwot et al.[2] demonstrated this through FCN-8s and DEM. Hybrid physics-DL models like SWE-GNN[1] integrate shallow water equations with graph neural networks, balancing accuracy and interpretability.

Despite progress, key issues remain in 3D mapping. 1) Limited 3D ground truth hinders model generalizability and forces reliance on synthetic hydrodynamic data. 2) Top vision models (e.g., SAM) are underutilized due to mismatches between the RGB inputs and geospatial data (multispectral imagery).

2 Research Questions

- What data augmentation strategies effectively mitigate flood depth label scarcity while preserving physical consistency?
- How can current advanced RGB-based visual models (such as SAM) be more efficiently applied to multichannel geospatial flood data?

3 Proposed Research Plan

2D Flood Mapping

- **Dual-branch Processing:**
  - RGB Branch: SAM generates coarse masks using text-derived prompts
  - Multichannel: CNN processes SAR/DEM data for multi-features
- **Cross-modal Fusion:**
  - Align SAM's mask proposals with CNN features via attention gates
  - Refine boundaries using DEM-derived slope constraints
- **Hybrid Prompt Generation:**
  - Use CLIP to convert text prompts into points prompts for SAM
  - Grounding DINO generates bounding boxes from text as SAM's prompts

3D Flood Depth Estimation

- **Physics-constrained augmentation:** Generate synthetic 3D labels using HEC-RAS simulations under varying rainfall scenarios
- **Graph-based interpolation:** Develop GNN architectures incorporating:
  - Terrain connectivity graphs from DEM derivatives (slope, aspect)
  - Sparse gauge measurements as node features
  - Mass conservation principles as edge constraints

Anticipated Contributions

- A multimodal framework for vision foundation models in flood analysis.
- Physics-guided data augmentation methodology addressing 3D label scarcity.
- Open-source flood mapping toolkit with interactive prompting capabilities.

4 Planed Timeline

- **Q1, Q2 2025:** Multimodal SAM adaptation and synthetic data generation
- **Q3, Q4 2025:** GNN model development and hybrid training
- **Q1, Q2 2026:** Cross-region validation and Comparative analysis
- **Q3, Q4 2026:** Tool chain integration and performance benchmarking

References

1. Bentivoglio, R., Isufi, E., Jonkman, S.N., Taormina, R.: Rapid spatio-temporal flood modelling via hydraulics-based graph neural networks. Hydrology and Earth System Sciences **27**(23), 4227–4246 (2023)
2. Gebrehiwot, A., Hashemi-Beni, L.: 3d inundation mapping: A comparison between deep learning image classification and geomorphic flood index approaches. Frontiers in Remote Sensing **3** (2022)
3. Karim, F., Armin, M.A., Ahmedt-Aristizabal, D., Tychsen-Smith, L., Petersson, L.: A review of hydrodynamic and machine learning approaches for flood inundation modeling. Water **15**(3) (2023). https://doi.org/10.3390/w15030566
4. Zhou, Y., Wu, W., Nathan, R., Wang, Q.J.: A rapid flood inundation modelling framework using deep learning with spatial reduction and reconstruction. Environmental Modelling Software **143** (2021)