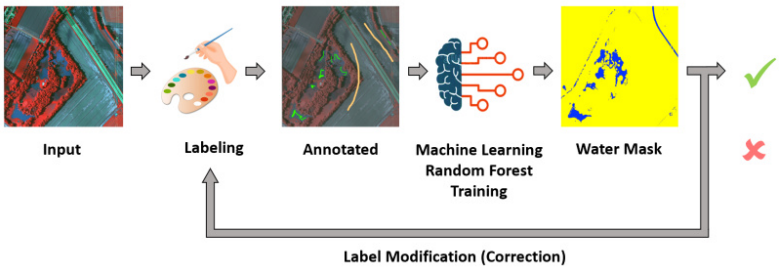
Hessisches Kompetenzzentrum
für Hochleistungsrechnen

Training of Deep Learning Models for Water Segmentation



Project Manager
Vladyslav Polushko

Principal Investigator
Prof. Dr. Thomas März

Project Term
2024 - 2025

Clusters
Lichtenberg II Cluster Darmstadt

Software
PyTorch

Additional Software
Albumentations, RasterIO, OpenCV , CUDA

JUSTUS-LIEBIG-UNIVERSITÄT GIESSEN
TECHNISCHE UNIVERSITÄT DARMSTADT
UNIKASSEL VERSITÄT
Philipps Universität Marburg

Introduction

Floods represent one of the most frequent and damaging natural disasters, causing loss of life, infrastructure failure and widespread economic disruption. Humanitarian organizations depend on rapid, accurate inundation maps to guide rescue operations, allocate resources and restore essential services. Yet manual analysis of large high-resolution image collections delays critical decisions. These challenges motivate automated image-analysis workflows.

Computer vision tools — especially Deep Learning (DL) segmentation models—automate water detection and reduce processing time. Because these DL-networks include millions of parameters and process large image volumes, they require substantial GPU compute and memory resources. Therefore, we leverage the Lichtenberg HPC Cluster to accelerate both training and inference, enabling deployment of resource-intensive segmentation models such as the transformer-based SegFormer-b5.

To address the shortage of annotated river-flood data, we introduced BlessemFlood21 [2], a 15 cm-resolution RGB orthomosaic of the July 2021 Erftstadt flood, annotated with pixel-precise water masks.

In our human-in-the-loop workflow, each tile is first rendered as an NIR-Blue-Red false-color composite to highlight flooded areas, leveraging the NIR band’s sensitivity to water. Annotators then mark sparse water and non-water strokes on these composites, guiding a Random Forest classifier to generate an initial mask. Finally, experts iteratively refine the mask until its boundaries align precisely with the visible floodwater.

We trained three state-of-the-art segmentation architectures—UNet++, DeepLabV3+ and SegFormer-b5—on the refined pairs of (RGB-image, mask) and evaluated them using Intersection-over-Union (IoU), Dice and Accuracy metrics.

To improve robustness under varied lighting, viewpoints and sensor artifacts, we tested nine augmentation groups and measure each group’s impact on performance. We assessed domain adaptability by sequentially fine-tuning our models: first on Blessem-Flood21 then on the coastal FloodNet [4] dataset, and then in reverse order. Both transfer directions yield notable IoU gains, showing that features learned on river-flood imagery can improve coastal segmentation and vice versa.

We demonstrate two applications: (1) estimating flood water level to decimeter precision by fitting predicted masks to a Digital Elevation Model (DEM) generated via structure-from-motion photogrammetry, and (2) detecting flooded roads by intersecting water masks with OpenStreetMap road data. Both methods underpin rapid, data-driven emergency response.

Methods

We surveyed the July 2021 Erftstadt flood using a gyrocopter-mounted PanX3 camera, capturing RGB+NIR imagery at 15 cm ground resolution. Images are georeferenced, mosaicked and split into uniform tiles for further processing.

Annotators draw sparse water/non-water strokes on NIR-RGB composites. A Random Forest proposes full water masks, which annotators refine until boundaries match visible inundation.

Although annotation leveraged NIR composites, all segmentation networks are trained on RGB inputs only, mirroring the typical drone sensor setup in flood-response operations. The tiles are divided into training, validation and test sets. We apply augmentations from the Albumentations library; geometric, photometric, blur/noise, optical distortions and weather effects on-the-fly to enhance robustness. On the Lichtenberg HPC Cluster, we trained the segmentation models UNet++, DeepLabV3+ and SegFormer-b5 using a binary cross-entropy loss and the Adam optimizer, and selected the best checkpoint by validation based on the IoU.

Each model was first fine-tuned on BlessemFlood21 and then on FloodNet and a second time in reversed order to assess the transfer performance via IoU, Dice and Accuracy metrics.

We estimated water depth by thresholding a 30 cm-resolution DEM at discrete elevation steps and selecting the level that maximizes IoU with the predicted mask. OpenStreetMap road semantics are rasterized and aligned to the orthomosaic of drone images via landmark-based polynomial warping. A pixel-wise intersection with predicted water masks then labels each road segment as flooded or passable.

Results

On the BlessemFlood21 test set, IoU spans roughly 75–91% across the evaluated architectures. DeepLabV3+ performs best, UNet++ follows, and SegFormer-b5 is lower. Dice and overall accuracy follow the same ranking and remain uniformly high.

For UNet++ under a fixed training budget, data augmentation improves IoU over the non-augmented baseline. Distortion provides the largest single-group gain (about 1.7 percentage points), with Blur/Noise next (about 0.9 percentage points). Photometric jitter and weather effects offer smaller but positive gains (each under 1 percentage point). Using the full set of augmentation groups yields the best overall improvement (about 2.0 percentage points over baseline).

Sequential fine-tuning from BlessemFlood21 to FloodNet increases IoU by approximately 59–60 percentage points across architectures.

In the reverse direction (FloodNet→BlessemFlood21), improvements are smaller but consistent, on the order of 0.4–1.5 percentage points.

Combining predicted flood masks with a high-resolution digital elevation model yields consistent water-level estimates with uncertainties of about 3–8 cm, indicating potential for rapid, practically useful flood-depth assessment.

Intersecting predicted flood extents with OpenStreetMap-derived road layers identifies submerged roads with precision and recall around 85%, suggesting utility for emergency route planning.

Discussion

The baseline segmentation results confirm that BlessemFlood21's high-resolution RGB imagery and precise human-in-the-loop annotations provide a solid foundation for training modern DL water segmentation models in river-flood scenarios.

The strong influence of Distortion and Blur/Noise augmentations account for drone-induced spatial warping and sensor blur. These findings align with prior studies emphasizing the need for models to resist common acquisition artifacts.

Cross-dataset transfer experiments reveal that sequential fine-tuning between river- and coastal-flood datasets can transfer knowledge effectively, despite domain gaps. This suggests potential for creating more generalizable flood mapping networks by leveraging diverse flood imagery.

Combining segmentation masks with a digital elevation model enables decimeter-level water-depth approximation, supporting preliminary flood-level assessment on site. Similarly, intersecting predicted masks with OpenStreetMap semantics offers a straightforward mechanism for flagging inundated road segments, which could aid rapid route planning in dynamic conditions.

Current limitations include manual mask corrections in areas with shadows or debris and reliance on hand-picked landmarks for geo-rectification, while evaluation is limited to a single river-flood event.

Our goals for future work are to automate annotation and alignment, to validate across multiple flood scenarios, and to integrate live Remote Sensing feeds for near-real-time inundation mapping.

Figures

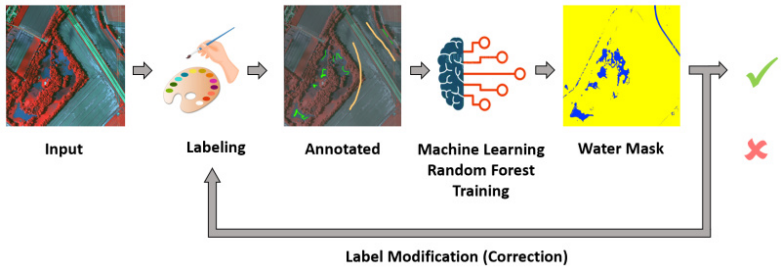


Figure 1: Human-in-the-loop Annotation Framework

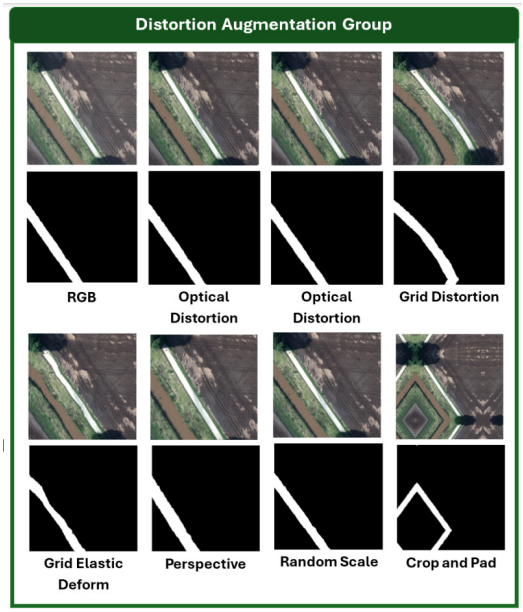
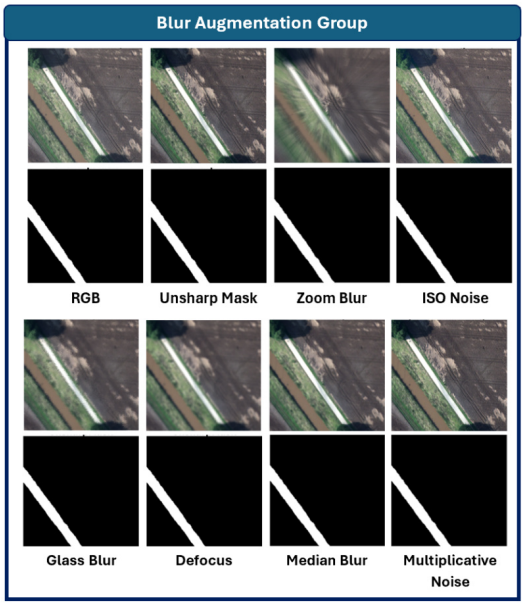


Figure 2: Distortion Augmentation group (best score improvements)



JUSTUS-LIEBIG-UNIVERSITÄT GIESSEN
TECHNISCHE UNIVERSITÄT DARMSTADT
UNIKASSEL VERSITÄT
Philipps Universität Marburg

Figure 3: Blur Augmentation group (second bessecond-bestt performance improvement)

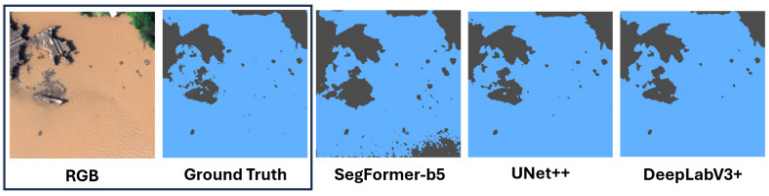


Figure 4: Qualitative water segmentation performance of the models SegFormer-b5, UNet++ and DeepLabV3+ when applied to the BlessemFlood21 dataset

Publications

[1] Polushko, V.; Hatic, D.; Rösch, R.; März, T.; Rauhut, M. and Weinmann, A.Remote sensing imagery for flood detection: Exploration of augmentation strategies (2025) https://arxiv.org/abs/2504.20203

[2] Polushko, V.; Jenal, A.; Bongartz, J.; Weber, I.; Hatic, D.; Rösch, R.; März, T.; Rauhut, M.; Weinmann, A. Blessemflood21: A high-resolution georeferenced dataset for advanced analysis of river flood scenarios. IEEE Access (2024). http://dx.doi.org/10.1109/ACCESS.2024.3487413

[3] Polushko, V.; Jenal, A.; Bongartz, J.; Weber, I.; Hatic, D.; Rösch, R.; März, T.; Rauhut, M.; Weinmann, A. Blessemflood21: Advancing flood analysis with a high-resolution georeferenced dataset for humanitarian aid support. In IEEE International Geoscience and Remote Sensing Symposium (IGARSS) (2024). http://dx.doi.org/10.1109/IGARSS53475.2024.10642128

Reference

[4] Rahnemoonfar, M.; Chowdhury, T.; Sarkar, A.; Varshney, D.; Yari, M.; Murphy, R. R. Floodnet: A high resolution aerial imagery dataset for post flood scene understanding. IEEE Access, 9:89644–89654 (2021). http://dx.doi.org/10.1109/ACCESS.2021.3090981

Last Update: 2025-08-25 18:24