Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 00 (2025) 000–000

Procedia Computer Science

www.elsevier.com/locate/procedia

29th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES 2025)

Beyond Major Floods: Deep Learning for Detecting Shallow Water Inundation in Agricultural Areas

Phongsakon Mark Konrad[a], Toygar Tanyel[b], Serkan Ayvaz[a,*]

[a]*Centre for Industrial Software, University of Southern Denmark, 6400 Sønderborg, Denmark*
[b]*Department of Electronics and Communication Engineering, Istanbul Technical University, 34469 Istanbul, Turkiye*

Abstract

Flood detection using satellite imagery is crucial for environmental monitoring and disaster management, especially in rural and agricultural regions where even minor water inundation can disrupt farmland accessibility and road safety. Sentinel-1 Synthetic Aperture Radar (SAR) imagery offers a robust solution for mapping water under various weather conditions. Although deep learning-based segmentation methods have shown promising results for flood detection, their comparative performance in agricultural landscapes, including small-scale surface water dynamics, remains underexplored. In this study, we introduced a three-class segmentation framework that distinguishes sea, inland water, and land, improving the flood detection accuracy in complex coastal farmland. Ten different deep learning models were evaluated for segmentation using Sentinel-1 VH polarization decibel values. We further investigated anomaly detection via autoencoders and variational autoencoders to track temporal changes in flood-prone areas. To handle large-scale satellite imagery more effectively, we tiled the images, ensuring that the segmentation models could process high-resolution data efficiently. The evaluations showed that the DeepLabv3+ and hybrid ResNet-UNet models outperformed the others. Despite having a more lightweight architecture and lower computational and memory resource requirements, the ResNet-UNet model achieved predictive performance comparable to that of DeepLabv3+.

Keywords: Sentinel-1; Satellite Imagery; Deep Learning; Segmentation; Anomaly Detection; Environmental Monitoring.

1. Introduction

Accurate flood detection is essential not only for managing large-scale disaster events but also for addressing smaller, localized occurrences of water inundation—particularly in agricultural areas where even minor flooding can disrupt farm operations. Synthetic Aperture Radar (SAR) data from Sentinel-1 has proven valuable for mapping water under various weather conditions, thanks to its cloud-penetrating capabilities [2, 31]. Existing methods for water

* Corresponding author. Address: Alsion 2, 6400 Sønderborg, Denmark
E-mail address: seay@mmmi.sdu.dk

detection range from straightforward threshold-based approaches [28, 40] to advanced deep learning techniques [8], each with varying strengths in accuracy and scalability.

However, micro-scale floods occurring in fields or on agricultural roads attract comparatively less attention than major floods. While official road networks are often monitored and maintained, farm tracks and rural paths can become unsafe or impassable due to water accumulation, threatening both personal safety and logistics [42]. In these settings, timely detection of even small waterlogging is key to preventing equipment damage, minimizing delays, and maintaining crop yields and overall farmland productivity.

Deep learning approaches have the potential to improve the detection of water bodies in these complex landscapes by capturing subtle patterns that traditional techniques may fail to recognize. However, to our knowledge, no prior studies have systematically evaluated multiple segmentation architectures for water inundation detection in agricultural settings. Existing research predominantly addresses urban flooding scenarios or broad-scale water segmentation tasks [3, 4], thereby leaving a gap in research for rural and farmland environments. Coastal agricultural zones, in particular, add an extra layer of complexity, blending sea, inland water, and land in close proximity.

In this study, we present a comprehensive evaluation of ten deep learning-based segmentation methods using Sentinel-1 VH polarization decibel values. Moving beyond the conventional two-class water–land categorization [20], we adopt a three-class segmentation framework that captures sea, inland water, and land. Additionally, we explore anomaly detection using statistical sliding-window method, autoencoders and variational autoencoders to monitor temporal changes in water extent. By employing an image-tiling strategy, we were able to process large volumes of high-resolution imagery efficiently.

As a use case, we applied our approach to the region of Southern Denmark, an area characterized by agricultural use and coastal influences. This application highlights the potential of SAR-based flood detection for improving road safety, protecting agricultural machinery, and ensuring farm accessibility. The overarching aim of this work is to provide actionable insights for stakeholders including farmers, local officials, and disaster management teams, about the effectiveness and scalability of deep learning approaches in capturing both large-scale flood events and localized, often-overlooked inundations that can significantly impact rural communities.

2. Related Work

Flood detection using Sentinel-1 SAR has been extensively studied through various methodologies. Comprehensive reviews on this topic, such as those by Amitrano et al. [2], Adeli et al. [1], and more recent perspectives on geospatial AI for flood mapping [24], provide an overview of key approaches, ranging from threshold-based techniques to deep learning-driven segmentation. These studies summarize the advantages and limitations of different methods, highlighting the effectiveness of SAR imagery in flood monitoring across diverse environments.

Early studies often relied on thresholding or basic classification and segmentation techniques, which provide efficient large-scale flood delineation but struggle with detecting subtle inundations. For instance, the near real-time flood detection framework by Martinis et al. [29] refined thresholding techniques by integrating segmentation-based classification for high-resolution SAR images. Meanwhile, Bartsch et al. [5] explored SAR-based monitoring of open water dynamics at high latitudes, focusing on long-term land surface modeling. Later, Martinis et al. [28] compared several operational approaches for SAR data, while Tran et al. [40] employed thresholding to map surface water in the Mekong Delta. Other threshold-based methods, such as those proposed by Bioresita et al. [7], demonstrated rapid and automatic water-mapping capabilities for emergency response.

Recent advancements in deep learning have shown immense promise in segmenting water in more complex scenarios. CNN-based methods [8, 14] and advanced architectures like vision transformers [37] leverage Sentinel-1 backscatter differences to distinguish flooded and non-flooded areas. The development of curated datasets, such as STURM-Flood [33], is further facilitating the training and benchmarking of such models. Ongoing research continues to yield architectural innovations, including hybrid CNN-Transformer models like Shuffle Window Transformer DeepLabV3+ [27] and DBSANet [26], aimed at improving segmentation accuracy and efficiency. Efforts are also intensifying to enhance the extraction of finer details and smaller water bodies, for example, by developing networks with advanced feature fusion and noise suppression techniques [13]. Fusion approaches integrating Sentinel-2 optical data [18, 39] also continue to refine flood delineation, particularly in vegetated regions.

The focus of flood detection is increasingly shifting towards more granular applications, including agricultural settings. For instance, Li et al. [25] recently utilized multiple deep learning models for the fine-scale identification of agricultural flooding disaster areas. While these advancements demonstrate significant progress in leveraging SAR data, and deep learning is increasingly applied to various agricultural flood scenarios [3, 25], a comprehensive comparative evaluation of a wide array of deep learning architectures specifically for detecting subtle, shallow water inundation—not necessarily indicative of large-scale disaster events in complex coastal agricultural landscapes remains less explored. The specific challenge of accurately distinguishing between sea, inland water, and land using a dedicated three-class segmentation framework to map minor inundations affecting farm accessibility and rural paths, and the systematic application of anomaly detection techniques to track their temporal dynamics, represent areas warranting further dedicated investigation. Addressing these underexplored scenarios by systematically evaluating ten distinct deep learning models for this nuanced task, focusing on Sentinel-1 VH polarization and a three-class framework, is crucial for improving safety, mitigating equipment damage, and maintaining farm operations, which serves as the primary motivation for our work.

3. Proposed Framework

We introduce a multistage pipeline for accurate flood mapping and temporal anomaly detection in agricultural landscapes. Unlike conventional approaches that only distinguish between water and land, the proposed framework employs a three-class framework separating *land*, *sea*, and *inland water* based on SAR image data. By integrating this detailed segmentation with temporal anomaly detection, the proposed pipeline identifies and tracks subtle water accumulations that might otherwise be overlooked. Figure 1 demonstrates an overview of the proposed architecture.

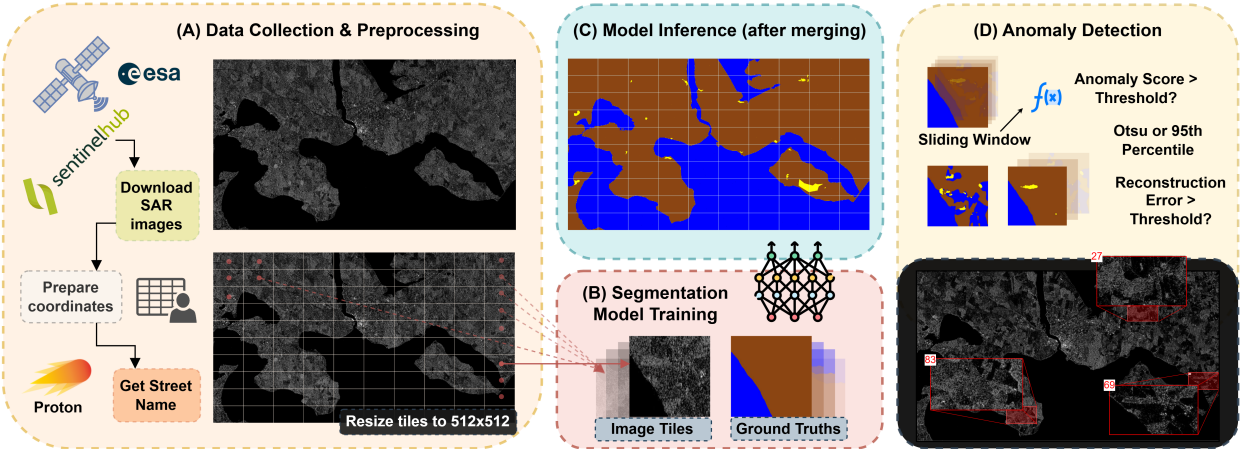

Fig. 1: Proposed pipeline for SAR image analysis and anomaly detection. (A) Data collection and preprocessing: SAR images are downloaded from Sentinel Hub, coordinates are determined, and images are tiled to 256x139 pixels based on resolution and the desired quantity of tiles. (B) Segmentation model training: Image tiles from one sample are used to train a model with corresponding manually labeled ground truth. (C) Model inference: The trained model generates segmented outputs at the tile level. (D) Anomaly detection: A sliding window approach with local mean intensity and pixel-level class ratios is used to compute anomaly scores with a threshold at the 95th percentile, or a reconstruction approach using an autoencoder (AE) or a variational autoencoder (VAE) with reconstruction errors thresholded using Otsu's method.

3.1. Data Collection and Augmentation

The first step in the pipeline involves acquiring Sentinel-1 SAR imagery in VH polarization of the region of interest (ROI), which spans from latitude 54.8339 ° to 55.0847 ° N and longitude 9.5856 ° to 10.0855 ° E. The images were then divided into 256 × 139 pixel tiles, creating a grid of 10 rows by 10 columns for a total of 100 tiles. Each tile was annotated with relevant metadata, including geographic coordinates and timestamps, to facilitate precise spatial-temporal tracking. The street names were also retrieved for each tile, if available, using reverse geocoding through

Photon, an open-source geocoder built for OpenStreetMap data, providing reverse geocoding services[1]. The meta data was saved in a CSV file. Later, for training, we resized the image to 512×512 pixel.

Due to the complexity of coastal agricultural landscapes and the presence of small dispersed water bodies, we manually annotated a single SAR image into three classes, *land*, *inland water*, and *sea* prior to the tiling procedure. To ensure accurate identification and prevent the overlooking of small inland water bodies, we cross-validated the manual annotations against high-resolution aerial images obtained from *Google Maps*. Annotations and augmentation were performed using the platform *Roboflow*, an end-to-end computer vision platform that simplifies data collection, annotation, model training, and deployment [2].

To improve the generalization of the model, we applied standard augmentation procedures such as random horizontal and vertical rotations, brightness modifications of approximately ±15%, and slight noise injection. After augmentation, each tile was normalized to an intensity range of [0, 1], preserving the essential details of radar backscatter that support the accurate detection of minor flood events. For supervised segmentation, the initial dataset from one non-anomalous day expanded from 100 original tiles to 300 through augmentations. For anomaly detection models, we retrieved Sentinel-1 SAR images of the year 2023, generating a larger dataset of 12,000 unlabeled tiles.

3.2. Segmentation Model Preprocessing and Hyperparameter Tuning

To ensure robust evaluation and prevent data leakage, we applied a two-stage data splitting procedure. Initially, we divide the dataset into training sets (60%), validation sets (20%), and tests sets (20%) for hyperparameter tuning using Bayesian optimization. After determining optimal hyperparameters, the training and validation sets (80% total) were combined and utilized in a 5-fold cross-validation approach. Within each fold, 80% of the combined set (64% of the original dataset) was used for training, and 20% (16% of the original dataset) for validation. The initial 20% test set remained untouched for final evaluation.

To address the class imbalance, particularly regarding the minority class *inland water*, all models were trained using *weighted cross-entropy* loss to improve predictive performance. Additionally, early stopping based on validation IoU was applied with a patience of 10 epochs to optimize training efficiency and mitigate overfitting.

Before the final training phase, we performed Bayesian optimization [34] using the Weights & Biases platform, an AI development platform that offers tools for tracking machine learning experiments, model management, and collaborative AI application development[3] to minimize validation loss. Key hyperparameters, including optimizer type (SGD, Adam, or AdamW), learning rate, and batch size, were systematically varied and evaluated to optimize convergence and generalization [32, 15]. The final models were trained using the best hyperparameters obtained through tuning.

3.3. Segmentation Models

We evaluated ten segmentation architectures, each selected for its capacity to handle varying complexities in satellite imagery. Classic encoder–decoder networks such as *UNet* [36], *UNet++* [44], *FC-DenseNet* [19], and *PSPNet* [43] serve as robust baselines. More advanced designs, including *ResNet-UNet* [16], *EfficientNet-UNet* [38], *DeepLabV3* [11], and *DeepLabV3+* [12], incorporate deeper convolutional backbones to capture subtle features more effectively. We also investigated transformer-based models such as *SegFormer* [41] and *Swin-UNet* [9], which excel at capturing both local and global spatial dependencies.

3.4. Anomaly Detection Models

We explored three distinct commonly used anomaly detection methods to identify subtle temporal deviations within satellite imagery.

[1] https://photon.komoot.io/
[2] https://roboflow.com/
[3] https://wandb.ai

3.4.1. Statistical Thresholding

We employed a sliding-window statistical approach, grouping tiles based on their indices from the prepared CSV file, followed by chronological sorting. Within each sliding window (size = 10), we computed the mean and standard deviation for the land ratio and mean intensity features. Anomalies are detected using dynamically computed percentile-based thresholds (e.g., 95th percentile) of the anomaly scores.

3.4.2. Autoencoder (AE)

An autoencoder [17] is trained on typical tile patterns to identify anomalous tiles based on elevated reconstruction errors. The Otsu method [35] is utilized to dynamically determine the optimal threshold for these errors.

3.4.3. Variational Autoencoder (VAE)

Extending the AE concept into a probabilistic domain, a variational autoencoder [21] encodes tile features into latent space distributions. Anomalies are detected based on a combination of reconstruction error and the Kullback–Leibler (KL) divergence [23], weighted by a tunable parameter β. The Otsu method is also applied here to optimize the threshold for reconstruction errors.

3.4.4. Anomaly Visualization

Anomalies identified by each method are visualized directly on satellite imagery using transparent bounding boxes based on tile metadata. These visualizations are further contextualized with concurrent weather data (e.g., temperature, humidity, precipitation) acquired via the *Open-Meteo API*, which offers global weather data services without API key requirements[4]. To validate the accuracy of our anomaly detection models in identifying environmental changes caused by extreme weather conditions, we utilized the weather data to identify past flooding events in the areas of interest as ground truth.

3.5. Experimental Setup

Experiments were performed on the Nvidia Jetson AGX Orin 64GB Developer Kit[5], leveraging Nvidia's Ampere architecture featuring 2048 CUDA cores and 64 Tensor cores. This hardware offers a computational capability of up to 275 TOPS (Tera Operations per Second) with configurable power settings ranging from 15W to 60W, making it suitable for efficient inference and evaluation tasks.

4. Results

To assess the effectiveness of the proposed approach, we performed several evaluations. For quantifying segmentation performance, we employed the Intersection over Union (IoU) and Dice coefficient, two prevalent metrics in image segmentation [30]. IoU, also known as the Jaccard Index, quantifies the overlap as the area of intersection divided by the area of union between the predicted and ground truth regions. The Dice coefficient (related to the F1-score) is calculated as twice the intersection divided by the sum of the areas of the predicted and ground truth regions; it often yields slightly higher scores than IoU and is considered by some to be more robust to class imbalance when dealing with smaller foreground objects. While both metrics measure similarity, IoU is generally considered a stricter evaluation. The segmentation evaluation results indicated that *ResNet-UNet* achieved the highest overall IoU (0.85 ± 0.05), whereas *DeepLabV3+* yielded the highest overall Dice (0.95 ± 0.02) as shown in Table 1. Multiple architectures, including these two, reached near-perfect metrics (0.98–0.99 IoU and about 1.00 Dice) for both *land* and *sea* classes, reflecting similar effectiveness in segmenting these larger, more distinct categories.

In contrast, *inland water* consistently yielded lower IoU and Dice across all methods, with best inland-water IoU values of 0.56 ± 0.14 (ResNet-UNet) and 0.54 ± 0.07 (DeepLabV3+). Transformer-based models such as *SwinUNet* and *SegFormer* also showed strong land and sea performance, though their inland-water IoU and Dice values remained below those of ResNet-UNet and DeepLabV3+. Other encoder-decoder architectures (e.g., *UNet*, *UNet++*,

[4] https://open-meteo.com/en/docs
[5] https://www.nvidia.com/en-us/autonomous-machines/embedded-systems/jetson-orin/

FC-DenseNet, and *EfficientNet-UNet*) demonstrated moderate or high segmentation accuracy for land and sea but exhibited reduced metrics for inland water, with some approaches displaying near-zero IoU in that class.

	Overall		Land		Sea		Inland Water	
Model	**IoU**	**Dice**	**IoU**	**Dice**	**IoU**	**Dice**	**IoU**	**Dice**
UNet	0.56 ± 0.22	0.59 ± 0.21	0.87 ± 0.21	0.92 ± 0.14	0.78 ± 0.42	0.80 ± 0.42	0.04 ± 0.09	0.07 ± 0.15
UNet++	0.65 ± 0.00	0.78 ± 0.00	0.98 ± 0.00	0.99 ± 0.00	0.97 ± 0.00	0.99 ± 0.00	0.00 ± 0.00	0.38 ± 0.00
FC-DenseNet	0.59 ± 0.01	0.83 ± 0.01	0.97 ± 0.00	0.99 ± 0.00	0.81 ± 0.02	0.84 ± 0.02	0.00 ± 0.00	0.67 ± 0.00
PSPNet	0.76 ± 0.08	0.85 ± 0.08	0.96 ± 0.05	0.98 ± 0.02	0.95 ± 0.07	0.97 ± 0.04	0.36 ± 0.14	0.59 ± 0.16
ResNet-UNet	**0.85 ± 0.05**	0.91 ± 0.06	**0.99 ± 0.00**	**1.00 ± 0.00**	**0.99 ± 0.00**	**1.00 ± 0.00**	**0.56 ± 0.14**	0.74 ± 0.17
EfficientNet-UNet	0.67 ± 0.07	0.80 ± 0.12	0.97 ± 0.04	0.98 ± 0.02	0.94 ± 0.09	0.96 ± 0.05	0.11 ± 0.11	0.46 ± 0.30
DeepLabV3	0.76 ± 0.04	0.81 ± 0.04	0.98 ± 0.01	0.99 ± 0.00	0.97 ± 0.01	0.99 ± 0.00	0.34 ± 0.12	0.44 ± 0.13
DeepLabV3+	0.84 ± 0.02	**0.95 ± 0.02**	**0.99 ± 0.00**	**1.00 ± 0.00**	**0.99 ± 0.00**	**1.00 ± 0.00**	0.54 ± 0.07	**0.86 ± 0.07**
SegFormer	0.77 ± 0.01	0.83 ± 0.05	0.98 ± 0.00	0.99 ± 0.00	0.98 ± 0.00	0.99 ± 0.00	0.36 ± 0.03	0.50 ± 0.14
SwinUNet	0.80 ± 0.02	0.86 ± 0.02	0.98 ± 0.00	0.99 ± 0.00	0.97 ± 0.00	0.99 ± 0.00	0.45 ± 0.06	0.59 ± 0.06

Table 1: Quantitative performance comparison of segmentation models on the testing dataset. The table reports Intersection over Union (IoU) and Dice similarity scores for overall performance and specific land cover classes: Land, Sea, and Inland Water. The best-performing values for each category are highlighted in bold.

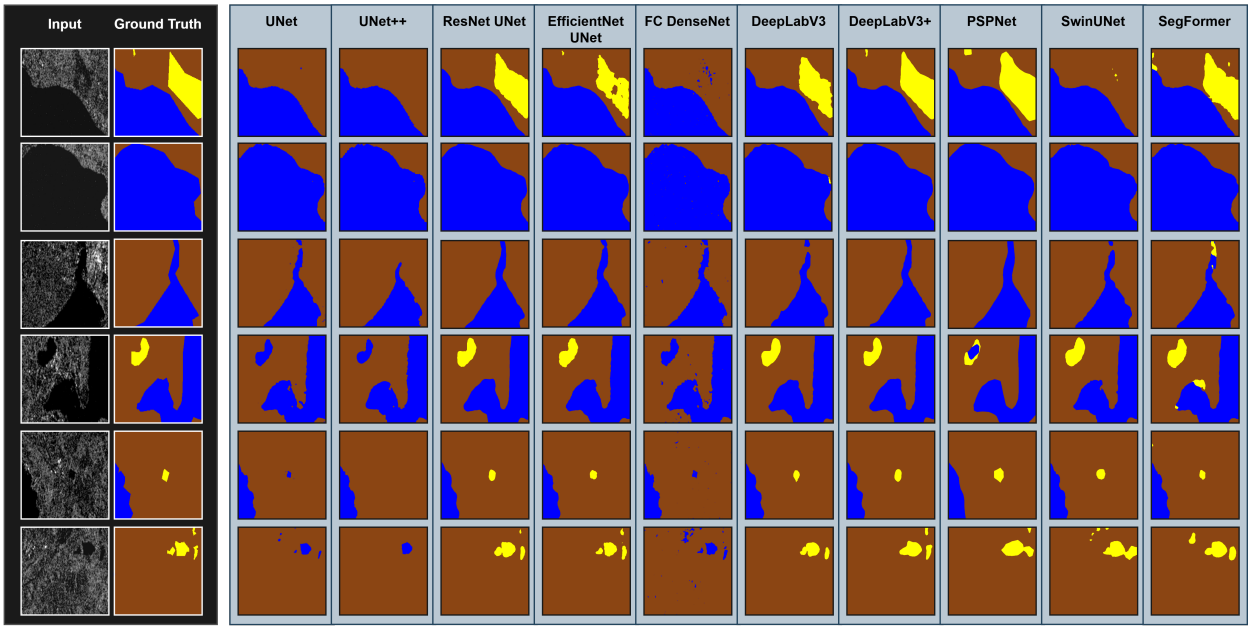

Fig. 2: Qualitative comparison of semantic segmentation models on SAR imagery. The figure presents input images, corresponding ground truth annotations, and segmentation predictions from various models. The models exhibit varying levels of accuracy in distinguishing different regions: land, sea, and inland water.

Additionally, we evaluated the computational resource utilization of the models using metrics including maximum GPU memory consumption (MB), inference throughput (images per second), throughput samples (samples per second), and inference time (seconds). The results showed that *SwinUNet* had the lowest maximum GPU memory requirement (2651.63 ± 160.33 MB), whereas *UNet* required the highest (23722.45 ± 161.10 MB). In terms of inference throughput, *ResNet-UNet* achieved the highest rate (7.57 ± 0.19 images per second) and also recorded the shortest inference time (7.93 ± 0.21 seconds). Other models generally resulted in a mid-range GPU memory footprint of about 3000–14000 MB, with inference speeds around 5–7 images per second. *FC DenseNet* exhibited the slowest throughput (2.71 ± 0.03 images per second) and the longest inference time (22.15 ± 0.28 seconds), while *UNet++*,

DeepLabV3, and *DeepLabV3+* showed intermediate resource demands and processing speeds. These values indicated notable variations in memory consumption and computational efficiency across all examined models.

Model	Max GPU Memory (MB) ↓	Inference Throughput (img/s) ↑	Throughput Samples (samples/s) ↑	Inference Time (s) ↓
UNet	23722.45 ± 161.10	5.64 ± 0.17	4.47 ± 0.02	10.65 ± 0.33
UNet++	12186.40 ± 160.42	5.54 ± 0.18	4.51 ± 0.01	10.84 ± 0.36
FC-DenseNet	14025.29 ± 17.92	2.71 ± 0.03	1.16 ± 0.00	22.15 ± 0.28
PSPNet	6477.77 ± 650.09	6.90 ± 0.17	10.85 ± 0.02	8.70 ± 0.21
ResNet-UNet	3071.44 ± 124.76	**7.57 ± 0.19**	**17.95 ± 0.10**	**7.93 ± 0.21**
EfficientNet-UNet	4631.39 ± 30.23	7.12 ± 0.20	12.23 ± 0.01	8.44 ± 0.24
DeepLabV3	9709.23 ± 211.40	5.64 ± 0.09	5.27 ± 0.00	10.65 ± 0.17
DeepLabV3+	9640.03 ± 211.46	7.33 ± 0.24	11.50 ± 0.15	8.19 ± 0.27
SegFormer	3397.03 ± 19.85	6.96 ± 0.21	11.51 ± 0.14	8.63 ± 0.27
SwinUNet	**2651.63 ± 160.33**	7.16 ± 0.29	14.44 ± 0.05	8.40 ± 0.35

Table 2: Inference performance comparison of semantic segmentation models. The table presents Max GPU Memory consumption (MB), Inference Throughput (images per second), Throughput Samples (samples per second), and Inference Time (seconds). Lower memory usage and inference time are preferable, while higher throughput indicates better efficiency.

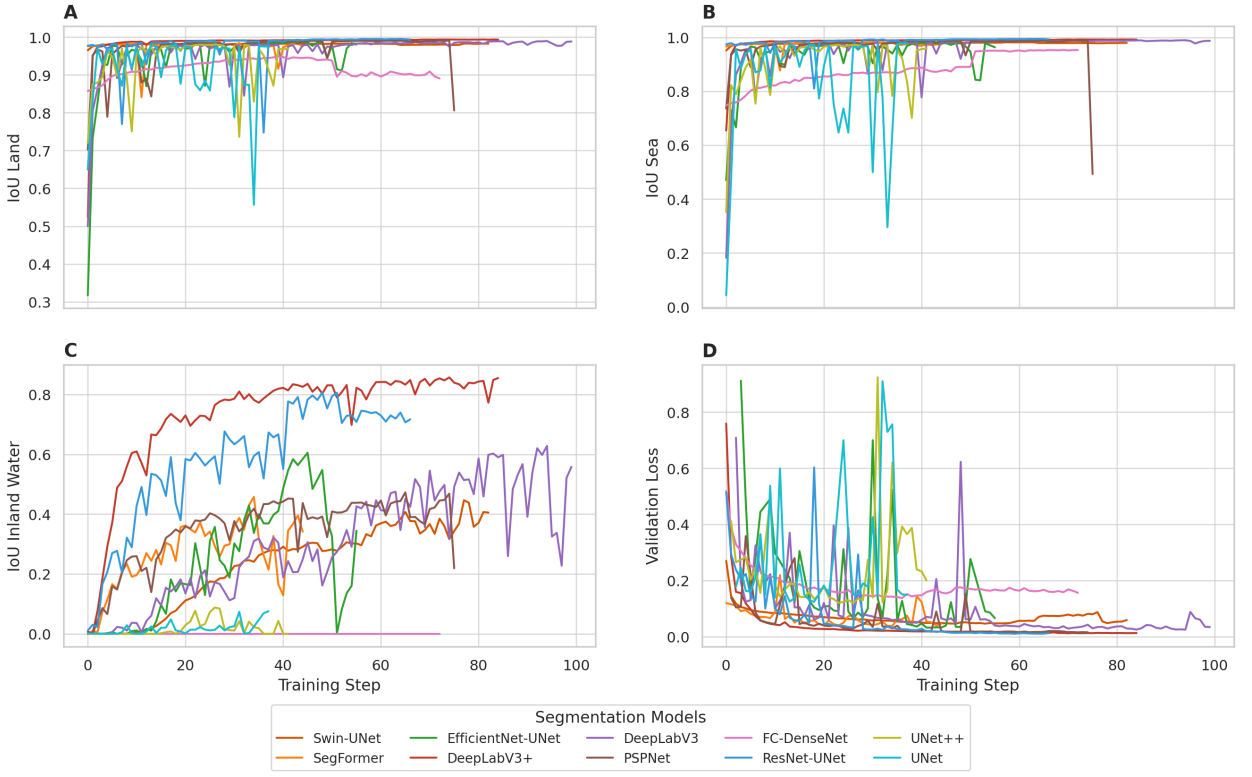

Fig. 3: Training performance of segmentation models over 100 training steps with macro-averaging of multiple folds. (A) Intersection over Union (IoU) for Land validation, (B) IoU for Sea validation, (C) IoU for Inland Water validation, and (D) Validation Loss. The models showed rapid convergence for Land and Sea classes, whereas Inland Water IoU exhibits more variation. Validation loss decreased over time, indicating overall learning progress, with some models stabilizing faster than others. Early stopping is evident for certain models, as their training curves plateau or stop updating before reaching 100 steps.

Qualitative comparison of semantic segmentation models on SAR imagery is illustrated in Figure 2. The figure displays input images, their corresponding ground truth annotations, and segmentation predictions of the evaluated

models. The models demonstrated varying degrees of accuracy in segmenting different classes. Some models were able to capture finer details more effectively, highlighting differences in segmentation performance across architectures.

Furthermore, we implemented early stopping with patience of 10 epochs and our main evaluation was the macro-average of the validation mean IoU. Figure 3 compares the performance and training stability of ten segmentation models by evaluating the Validation IoU (panel A,B,C) and the overall Validation Loss (panel D). Therefore, in Figure 3, the differences in the training steps are clearly visible. Among the observed architectures, **DeepLabV3+** and **ResNet-UNet** clearly stand out, achieving IoU scores around 0.8 and demonstrating smooth and stable convergence. In contrast, other models such as *EfficientNet-UNet*, *PSPNet*, *UNet++* consistently showed fluctuations during training.

5. Discussion

In this study, we introduced an end-to-end architecture for detecting abrupt changes in water inundation based on Earth observation data, framed as an anomaly detection problem. In addition, we evaluated the effectiveness of several state-of-the-art deep learning algorithms in identifying water inundation through satellite radar imagery. Due to cloud cover obstructing optical imaging on overcast days, we only considered using satellite radar imagery in the assessments.

More specifically, we focused on evaluating the ability of these models to detect small water accumulations within terrestrial environments. The detection of such small water bodies is both challenging and crucial, as their presence in agricultural areas can pose substantial risks to the safe operation of agricultural vehicles and adversely affect crop productivity.

With climate change, accurate detection of water accumulations will become even more important, as the frequency and duration of small-scale water inundation in land are expected to increase due to shifts in precipitation patterns, rising temperatures, and altered hydrological processes [22, 6]. More intense and frequent extreme rainfall events, driven by a warming climate, can lead to localized flooding, water pooling and reduced infiltration capacity, causing excess water to remain on the surface longer periods, particularly in low-lying agricultural fields and wetlands [10].

A noteworthy observation from the evaluations was that while most deep learning models effectively distinguished between land and sea segments, they encountered difficulties in accurately identifying inland water inundation. This challenge arises due to varying depths of water bodies, which are generally shallower than the sea and often share visual features more similar with land. On the other hand, when a tile contains a large water body, even if it is inland water, the models tend to exhibit a bias toward classifying it as sea during segmentation. However, this misclassification is not a major concern for anomaly detection models, as they primarily focus on tracking the changes over time.

Upon evaluating the performance of various deep learning models on satellite imagery, we observed differences in their ability to detect lakes, ponds, and small water bodies on land. Among the tested models, DeepLabv3+ and hybrid ResNet-UNet models outperformed the others. While DeepLabv3+ demonstrated high prediction accuracy, it required substantial memory and computational resources. On the other hand, ResNet-UNet exhibited comparable predictive performance despite its more lightweight architecture and lower resource demands. These findings suggest that the ResNet-UNet model offers a more efficient and practical solution for this use case and similar applications.

Furthermore, we assessed the performance of both weighted and unweighted ensemble approaches, while we observed a slight prediction improvement in the land and sea segmentation, the performance of ensemble approaches was worse in inland water segmentation compared to DeepLabv3+ and ResNet-UNet. This discrepancy occurred because these two constituent deep learning models significantly outperformed the remaining models in the evaluations.

A potential limitation of the proposed approach is its applicability to real-time decision-making for vehicles operating in agricultural fields, as the frequency of satellite data acquisition for the same location is relatively low. However, the proposed approach can be integrated with models that utilize edge computing and onboard sensor data to enable near real-time decision-making. Future work will focus on extending this methodology by integrating complementary edge-computing models through multimodal data fusion. Additionally, the dataset will be expanded to include diverse landscapes, aiming to improve the models' generalizability.

6. Conclusion

Detection of sudden changes in water accumulation in rural and agricultural areas is essential for maintaining the operational safety and accessibility for both vehicles and individuals. In this study, we proposed a framework for detecting changes in water inundation using satellite radar image data from Sentinel-1. To evaluate its effectiveness, we assessed the performances of ten deep learning models for image segmentation, introducing a three-class segmentation approach to enhance detection accuracy on complex coastal farmland. To facilitate large-scale flood detection, we tiled high-resolution satellite imagery for efficient processing. Our findings indicate that DeepLabv3+ and ResNet-UNet outperformed other models, with ResNet-UNet offering a balance between computational efficiency and predictive accuracy, whereas DeepLabv3+ required higher computational and memory resources for a similar performance. Additionally, anomaly detection through autoencoders and variational autoencoders effectively captured significant temporal changes in flood-prone regions.

Acknowledgements

This work is supported by the Innovation Fund Denmark for the project DIREC (9142-00001B). Additional support was provided by Danish Data Science Academy (DDSA) Visit Grant. The content is solely the responsibility of the authors and does not represent the official views of funding sources.

References

[1] Adeli, S., Salehi, B., Mahdianpari, M., Quackenbush, L.J., Brisco, B., Tamiminia, H., Shaw, S., 2020. Wetland monitoring using SAR data: a meta-analysis and comprehensive review. Remote Sensing 12, 2190.

[2] Amitrano, D., Di Martino, G., Di Simone, A., Imperatore, P., 2024. Flood detection with SAR: a review of techniques and datasets. Remote Sensing 16, 656.

[3] Bahrami, B., Arbabkhah, H., 2024. Enhanced flood detection through precise water segmentation using advanced deep learning models. Journal of Civil Engineering Researchers 6, 1–8.

[4] Bai, Y., Wu, W., Yang, Z., Yu, J., Zhao, B., Liu, X., Yang, H., Mas, E., Koshimura, S., 2021. Enhancement of detecting permanent water and temporary water in flood disasters by fusing Sentinel-1 and Sentinel-2 imagery using deep learning algorithms: demonstration of Sen1Floods11 benchmark datasets. Remote Sensing 13, 2220.

[5] Bartsch, A., Trofaier, A., Hayman, G., Sabel, D., Schlaffer, S., Clark, D., Blyth, E., 2012. Detection of open water dynamics with ENVISAT ASAR in support of land surface modelling at high latitudes. Biogeosciences 9, 703–714.

[6] Bevacqua, E., Rakovec, O., Schumacher, D.L., Kumar, R., Thober, S., Samaniego, L., Seneviratne, S.I., Zscheischler, J., 2024. Direct and lagged climate change effects intensified the 2022 european drought. Nature Geoscience , 1–8.

[7] Bioresita, F., Puissant, A., Stumpf, A., Malet, J.P., 2018. A method for automatic and rapid mapping of water surfaces from Sentinel-1 imagery. Remote Sensing 10, 217.

[8] Bonafilia, D., Tellman, B., Anderson, T., Issenberg, E., 2020. Sen1Floods11: a georeferenced dataset to train and test deep learning flood algorithms for Sentinel-1, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, pp. 210–211.

[9] Cao, H., Wang, Y., Chen, J., Jiang, D., Zhang, X., Tian, Q., Wang, M., 2021. Swin-Unet: Unet-like pure transformer for medical image segmentation. arXiv preprint arXiv:2105.05537 Later published in Expert Systems With Applications, 2024.

[10] Chang, H., Bonnette, M.R., 2016. Climate change and water-related ecosystem services: impacts of drought in California, USA. Ecosystem Health and Sustainability 2, e01254.

[11] Chen, L.C., Papandreou, G., Schroff, F., Adam, H., 2017. Rethinking atrous convolution for semantic image segmentation. arXiv preprint arXiv:1706.05587 .

[12] Chen, L.C., Zhu, Y., Papandreou, G., Schroff, F., Adam, H., 2018. Encoder-decoder with atrous separable convolution for semantic image segmentation, in: Proceedings of the European Conference on Computer Vision (ECCV), Springer. pp. 801–818. ArXiv:1802.02611.

[13] Gao, M., Dong, W., Chen, L., Wu, Z., 2025. Automatic extraction of water body from SAR images considering enhanced feature fusion and noise suppression. Applied Sciences 15, 2366. Published 22 February 2025.

[14] Ghosh, B., Garg, S., Motagh, M., Martinis, S., 2024. Automatic flood detection from Sentinel-1 data using a nested UNet model and a NASA benchmark dataset. PFG–Journal of Photogrammetry, Remote Sensing and Geoinformation Science 92, 1–18.

[15] Golmant, N., Vemuri, N., Yao, Z., Feinberg, V., Gholami, A., Rothauge, K., Mahoney, M.W., Gonzalez, J., 2018. On the computational inefficiency of large batch sizes for stochastic gradient descent. arXiv preprint arXiv:1811.12941 .

[16] He, K., Zhang, X., Ren, S., Sun, J., 2015. Deep residual learning for image recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778. ArXiv:1512.03385. Conference version published in 2016.

[17] Hinton, G.E., Salakhutdinov, R.R., 2006. Reducing the dimensionality of data with neural networks. Science 313, 504–507.

[18] Huang, M., Jin, S., 2020. Rapid flood mapping and evaluation with a supervised classifier and change detection in Shouguang using Sentinel-1 SAR and Sentinel-2 optical data. Remote Sensing 12, 2073.

[19] Jégou, S., Drozdzal, M., Vazquez, D., Romero, A., Bengio, Y., 2016. The one hundred layers Tiramisu: fully convolutional DenseNets for semantic segmentation, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 11–19. ArXiv:1611.09326. Conference version published in 2017.

[20] Kim, M.U., Oh, H., Lee, S.J., Choi, Y., Han, S., 2021. A large-scale dataset for water segmentation of SAR satellite, in: 2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), IEEE. pp. 9796–9801.

[21] Kingma, D.P., Welling, M., 2013. Auto-encoding variational Bayes. arXiv preprint arXiv:1312.6114 Conference version published at ICLR 2014.

[22] Konapala, G., Mishra, A.K., Wada, Y., Mann, M.E., 2020. Climate change will affect global water availability through compounding changes in seasonal precipitation and evaporation. Nature Communications 11, 3044.

[23] Kullback, S., Leibler, R.A., 1951. On information and sufficiency. Annals of Mathematical Statistics 22, 79–86.

[24] Lee, H., Li, W., 2025. Geospatial artificial intelligence for satellite-based flood extent mapping: concepts, advances, and future perspectives.

[25] Li, J., Gao, J., Chen, H., Shen, X., Zhu, X., Qiao, Y., 2025a. Fine-scale identification of agricultural flooding disaster areas based on Sentinel-1/2: a case study of Shengzhou, Zhejiang Province, China. Atmosphere 16, 420. Published 21 March 2025.

[26] Li, Y., Zhu, W., Wu, J., Zhang, R., Xu, X., Zhou, Y., 2025b. DBSANet: a dual-branch semantic aggregation network integrating CNNs and transformers for landslide detection in remote sensing images. Remote Sensing 17, 807. Published 22 February 2025.

[27] Liu, C., Ju, Z., Gao, X., Geng, A., 2024. Shuffle window transformer DeepLabV3+: a lightweight convolutional neural network and transformer based hybrid semantic segmentation network. Multimedia Tools and Applications Published online 06 May 2024.

[28] Martinis, S., Kuenzer, C., Wendleder, A., Huth, J., Twele, A., Roth, A., Dech, S., 2015. Comparing four operational SAR-based water and flood detection approaches. International Journal of Remote Sensing 36, 3519–3543.

[29] Martinis, S., Twele, A., Voigt, S., 2009. Towards operational near real-time flood detection using a split-based automatic thresholding procedure on high resolution TerraSAR-X data. Natural Hazards and Earth System Sciences 9, 303–314.

[30] Minaee, S., Boykov, Y.Y., Porikli, F., Plaza, A.J., Kehtarnavaz, N., Terzopoulos, D., 2020. Image segmentation using deep learning: a survey. IEEE Transactions on Pattern Analysis and Machine Intelligence 44, 3523–3542. ArXiv:2001.05566. Early access 2021. Final publication July 2022.

[31] Moreira, A., Prats-Iraola, P., Younis, M., Krieger, G., Hajnsek, I., Papathanassiou, K.P., 2013. A tutorial on synthetic aperture radar. IEEE Geoscience and Remote Sensing Magazine 1, 6–43.

[32] Mortazi, A., Cicek, V., Keles, E., Bagci, U., 2023. Selecting the best optimizers for deep learning based medical image segmentation. arXiv preprint arXiv:2307.13678 .

[33] Notarangelo, N.M., Wirion, C., van Winsen, F., Schnebele, E., Risser, L., Dugar, S.K., Kulkarni, H.K., Pinto, N., Jha, R.K., Fomferra, S.J., Stier, T., Navacchi, L., Ge, J., Appel, F., Rossi, L., 2025. STURM-Flood: a curated dataset for deep learning-based flood extent mapping leveraging Sentinel-1 and Sentinel-2 imagery. Big Earth Data Published online 24 Jan 2025.

[34] Onorato, G.M., Sovrano, F., Zese, R., Riguzzi, F., 2024. Bayesian optimization for hyperparameters tuning in neural networks, in: Italian Workshop on Machine Learning and Data Mining (WOA), pp. 273–280.

[35] Otsu, N., 1979. A threshold selection method from gray-level histograms. IEEE Transactions on Systems, Man, and Cybernetics 9, 62–66.

[36] Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: convolutional networks for biomedical image segmentation, in: Medical Image Computing and Computer-Assisted Intervention (MICCAI), Springer. pp. 234–241. ArXiv:1505.04597.

[37] Saleh, T., Weng, X., Holail, S., Hao, C., Xia, G.S., 2024. DAM-Net: flood detection from SAR imagery using differential attention metric-based vision transformers. ISPRS Journal of Photogrammetry and Remote Sensing 212, 440–453.

[38] Tan, M., Le, Q., 2019. EfficientNet: rethinking model scaling for convolutional neural networks, in: Proceedings of the 36th International Conference on Machine Learning, PMLR. pp. 6105–6114. ArXiv:1905.11946.

[39] Tavus, B., Kocaman, S., Nefeslioglu, H., Gokceoglu, C., 2020. A fusion approach for flood mapping using Sentinel-1 and Sentinel-2 datasets. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 43, 641–648.

[40] Tran, K.H., Menenti, M., Jia, L., 2022. Surface water mapping and flood monitoring in the Mekong Delta using Sentinel-1 SAR time series and Otsu threshold. Remote Sensing 14, 5721.

[41] Xie, E., Wang, W., Yu, Z., Anandkumar, A., Alvarez, J.M., Luo, P., 2021. SegFormer: simple and efficient design for semantic segmentation with transformers. Advances in Neural Information Processing Systems 34, 12077–12090. ArXiv:2105.15203.

[42] Zeedyk, B., 2006. Water harvesting from low-standard rural roads. Santa Fe, NM: Quivira Coalition .

[43] Zhao, H., Shi, J., Qi, X., Wang, X., Jia, J., 2016. Pyramid scene parsing network, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2881–2890. ArXiv:1612.01105. Conference version published in 2017.

[44] Zhou, Z., Rahman Siddiquee, M.M., Tajbakhsh, N., Liang, J., 2018. UNet++: a nested U-Net architecture for medical image segmentation, in: Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support, Springer. pp. 3–11. ArXiv:1807.10165.