

Water Surface Segmentation Results of Different Models

We selected two recent semantic segmentation models mentioned in the survey (Minaee et al., 2021), namely DeepLabV3+ (Chen et al., 2018) and HRNet (Sun et al., 2019). In addition, we selected another three recent semantic segmentation models, namely SegFormer (Xie et al., 2021), Mask2Former (Cheng et al., 2022) and SAN (Xu et al., 2023).

SegNet and the above recent models were trained, validated and tested on our datasets, and the results are shown in Table I.

Table I. The performance of the segmentation models on our training dataset, validation dataset and testing dataset (i.e., evaluation dataset), in terms of mean Pixel Accuracy (mPA) and mean Intersection over Union (mIoU).

Model	Training dataset		Validation dataset		Testing dataset	
	mPA (%)	mIoU (%)	mPA (%)	mIoU (%)	mPA (%)	mIoU (%)
SegNet	99.15	94.11	96.80	94.45	93.31	87.65
DeepLabV3+	99.10	95.21	97.14	94.35	95.49	91.54
HRNet	99.37	95.64	97.03	94.09	94.76	90.24
SegFormer	99.64	96.19	97.60	95.12	96.73	93.82
Mask2Former	98.74	96.25	98.08	96.05	97.20	94.71
SAN	99.53	99.09	98.58	97.33	97.05	94.24

Judging from the overall performance of each model on our three datasets, these recent models are better than SegNet.

In addition, we evaluated the performance of SegNet and the recent segmentation models on the three additional datasets mentioned in Wang et al. (2024). The three dataset are the WebCOOS (Wang et al., 2024), Deepflood (Chaudhary et al., 2020) and Sazara (Sazara et al., 2019). The performance of these models are shown in Table II. It should be noted that SegNet and the recent models were trained on our training dataset.

Table II. The performance of different segmentation models on various datasets, in terms of mPA and mIoU.

Model	Dataset	mPA (%)	mIoU (%)
SegNet	WebCOOS	68.49	50.39
	Deepflood	91.39	83.53
	Sazara	93.69	88.12
DeepLabV3+	WebCOOS	67.70	49.69
	Deepflood	93.00	86.31
	Sazara	94.95	90.38

HRNet	WebCOOS	71.15	52.48
	Deepflood	92.72	85.78
	Sazara	94.70	89.94
SegFormer	WebCOOS	71.08	51.91
	Deepflood	93.92	87.70
	Sazara	95.53	91.44
Mask2Former	WebCOOS	56.77	39.57
	Deepflood	95.21	90.01
	Sazara	97.24	94.67
SAN	WebCOOS	59.69	42.19
	Deepflood	93.33	87.29
	Sazara	95.09	90.52

SegNet and the recent models achieved high performance on the Sazara and Deepflood datasets, but lower performance on the WebCOOS dataset. The reason is that our training dataset has similar data distribution as the Deepflood and Sazara datasets, but there is a significant difference in data distribution between our training dataset and the WebCOOS dataset (our training dataset consists of Internet images, while the WebCOOS dataset consists of real-world surveillance camera images). In subsequent research, we need to supplement real-world surveillance camera images to improve the performance and generalization ability of the water surface segmentation models.

On the WebCOOS dataset, the performance of DeepLabV3+, Mask2Former, and SAN is lower than that of SegNet, while the performance of HRNet and SegFormer is higher than that of SegNet. On the Deepflood and Sazara datasets, all the recent models perform better than SegNet.

Only HRNet and SegFormer perform better than SegNet on the three datasets. In terms of mPA and mIoU on the three datasets, the overall performance of SegFormer is better than HRNet. Therefore, we finally chose to use SegFormer as the water surface segmentation model.

By using SegFormer, the water surface segmentation performance on the evaluation dataset was improved, with mPA increasing from 93.31% to 96.73%, and mIoU increasing from 87.65% to 93.82%. The identification accuracy of pedestrian submerged part was improved from 90.16% to 90.71%.