Table 1. Performances and limitations of the different fusion approaches

Fusion approach	Performances	Limitations
Prior-level	-Direct use of semantic information from images -Fast convergence -Low loss function -High classification accuracy.	-Problems of non-overlapping regions and uncertainties -Bit long process
Point-level	-Fast drive -Easy handling -No prior information is required.	- High cost - Not able to classify diversified urban contexts - Relatively low classification accuracy
Feature-level	-Objective data compression -Retaining enough important information	-Training loss higher -Features may not reflect the real objects.
Decision-level	-Non-interference of the two semantic segmentation processes -Good flexibility -Low-complexity -Learning the representation of independent features is allowed	-Impacted by the shortcomings of both classifiers. - Additional parameters for layers are required - More memory requirement

Table 2. Quantitative results of the proposed scenarios and the reference approach obtained by KPConv

urban scenes	Processes	F1-score	Recall	Precision	IoU
Scene 1	Baseline approach	0.68	0.74	0.68	0.58
	S1	**0.81**	**0.84**	**0.81**	**0.72**
	S2	0.70	0.76	0.70	0.60
Scene 2	Baseline approach	0.76	0.81	0.78	0.67
	S1	**0.87**	**0.89**	**0.87**	**0.81**
	S2	0.80	0.84	0.82	0.73

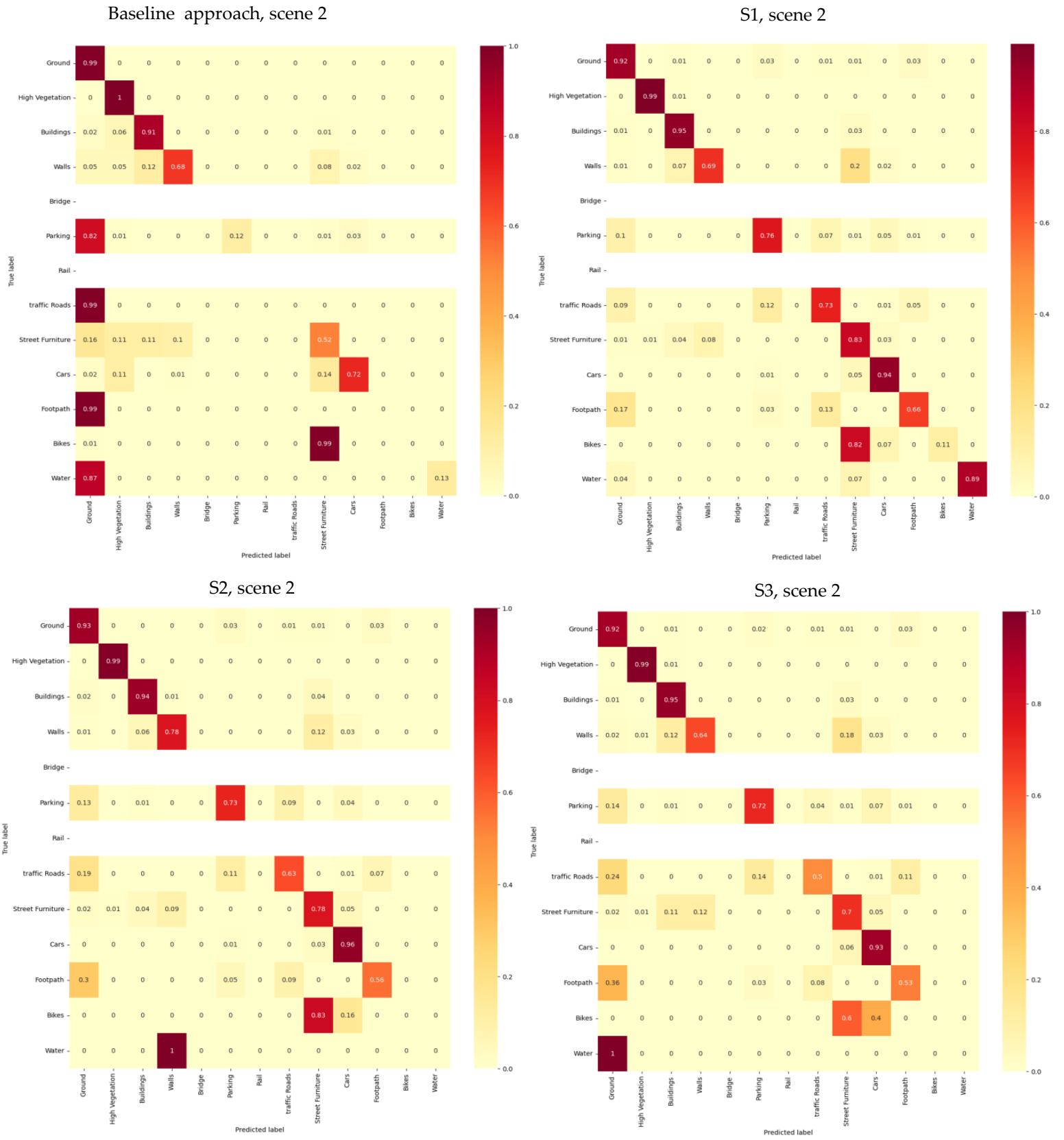


Figure 1. Normalized Confusion matrix of evaluated semantic segmentation approaches over the scene 2

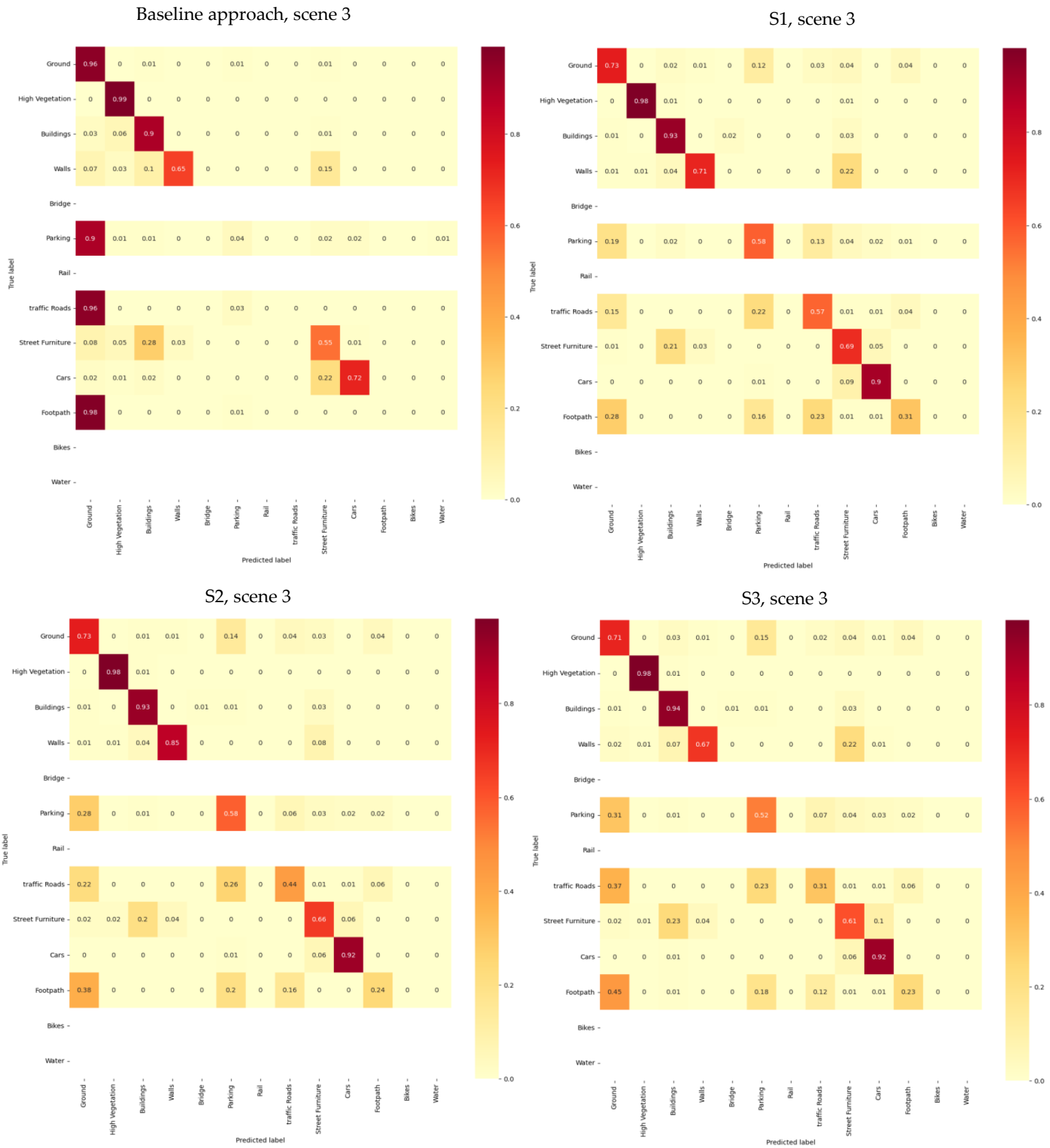
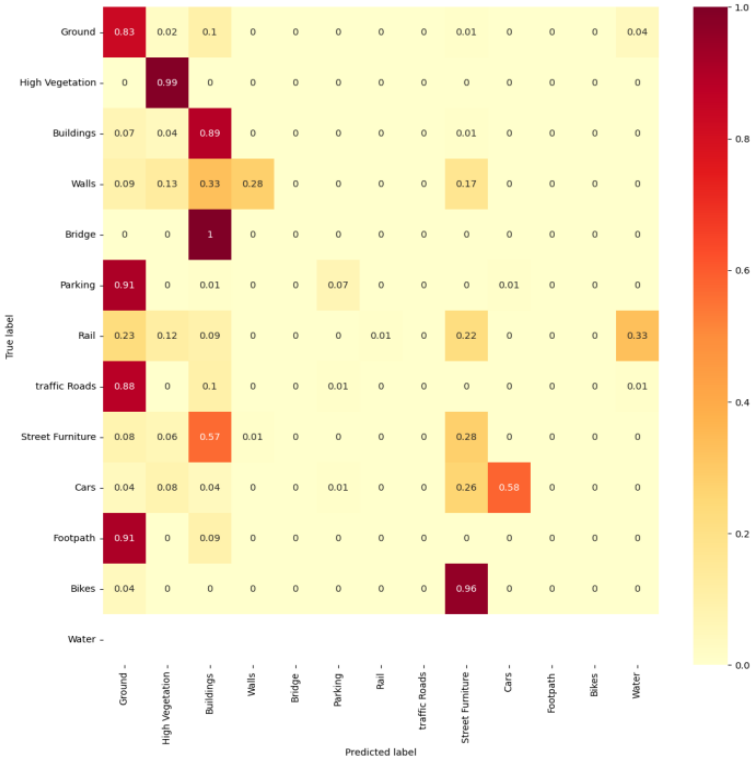
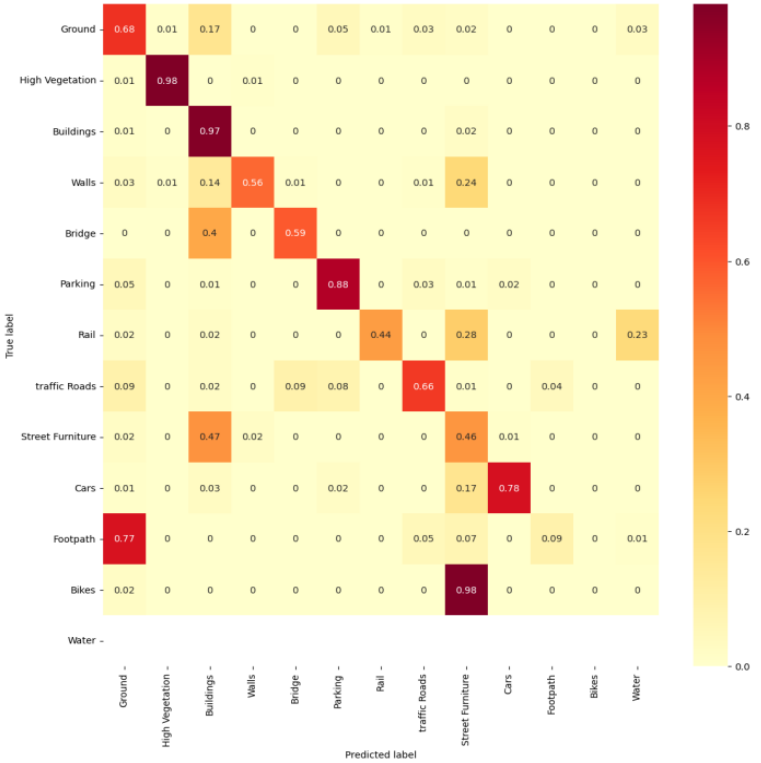


Figure 2. Normalized Confusion matrix of evaluated semantic segmentation approaches over the scene 3

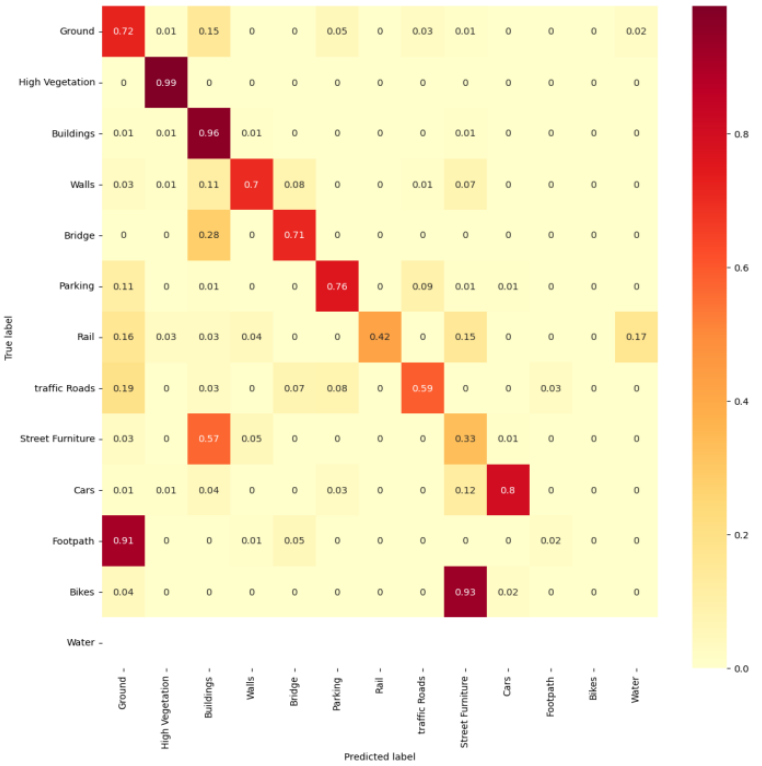
Baseline approach, scene 4



S1, scene 4



S2, scene 4



S3, scene 4

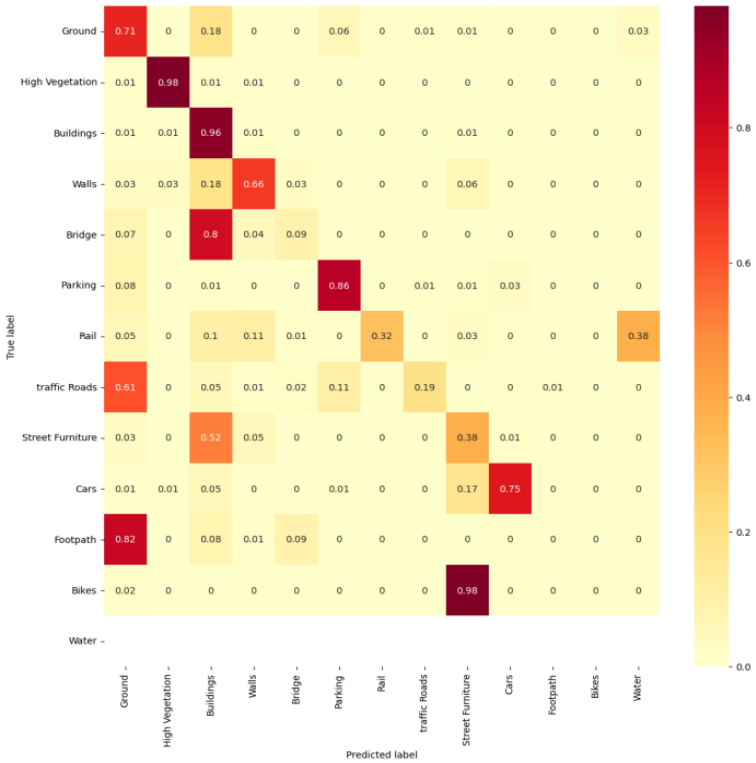


Figure 3. Normalized Confusion matrix of evaluated semantic segmentation approaches over the scene 4