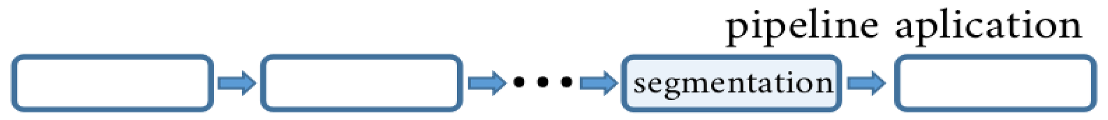
Graph Neural Blocks on Segmentation

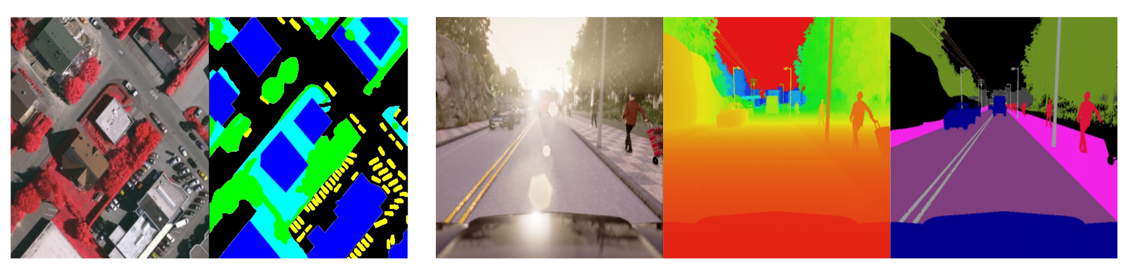
DARWIN.PILCO@IC.UNICAMP.BR AND ANTOINE.TABBONE@UNIV-LORRAINE.FR
Institute of Computing, University of Campinas, LORIA, Université de Lorraine

Context

Semantic Segmentation is one of the key steps in computer vision applications.



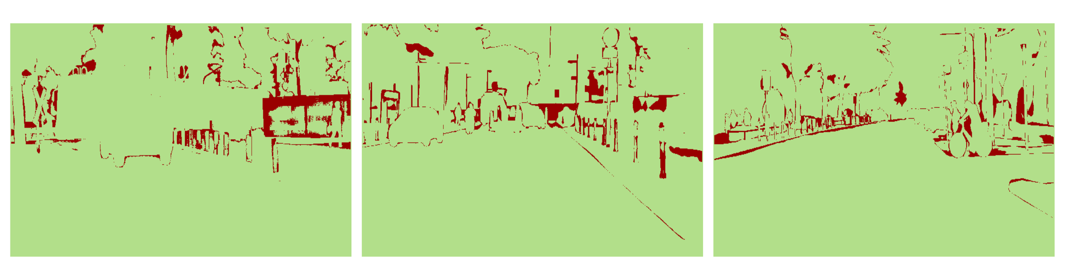
(a) X-ray (b) Agriculture



(c) Remote Sensing (d) Autonomous Driving

Semantic segmentation problems

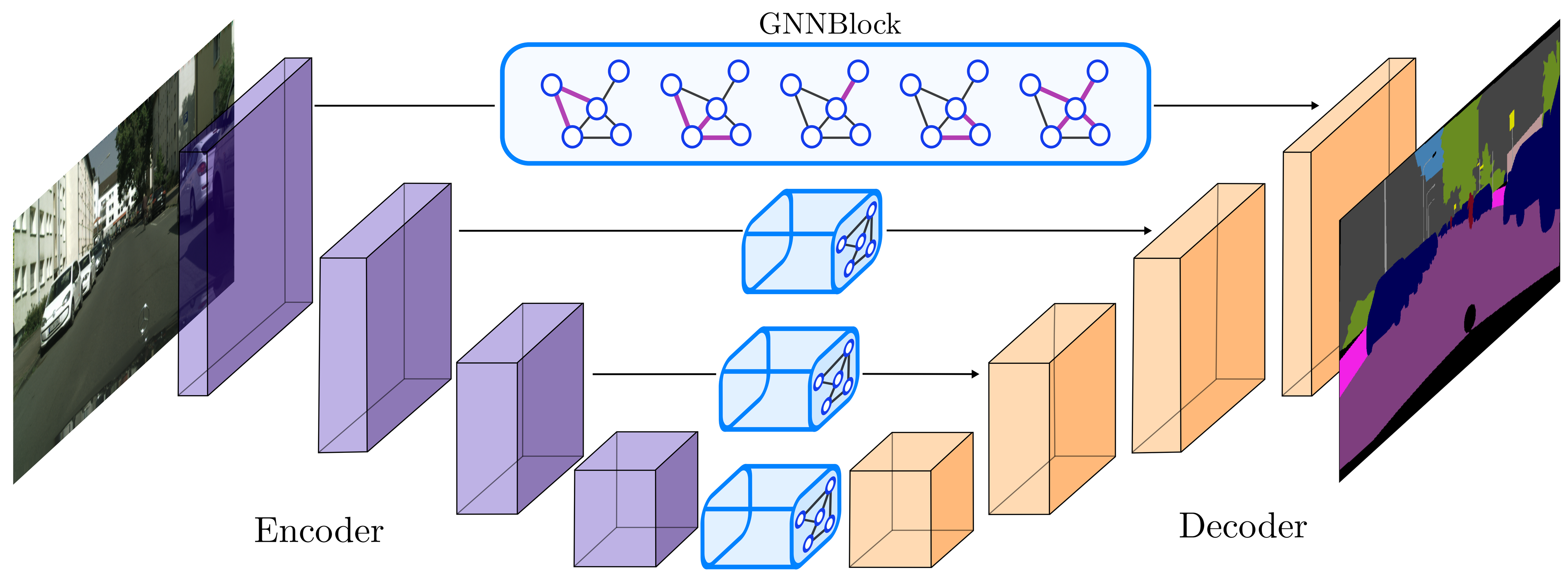
- Low-resolution in output heatmaps (solved)
- Loss of spatial precision (**remain**)

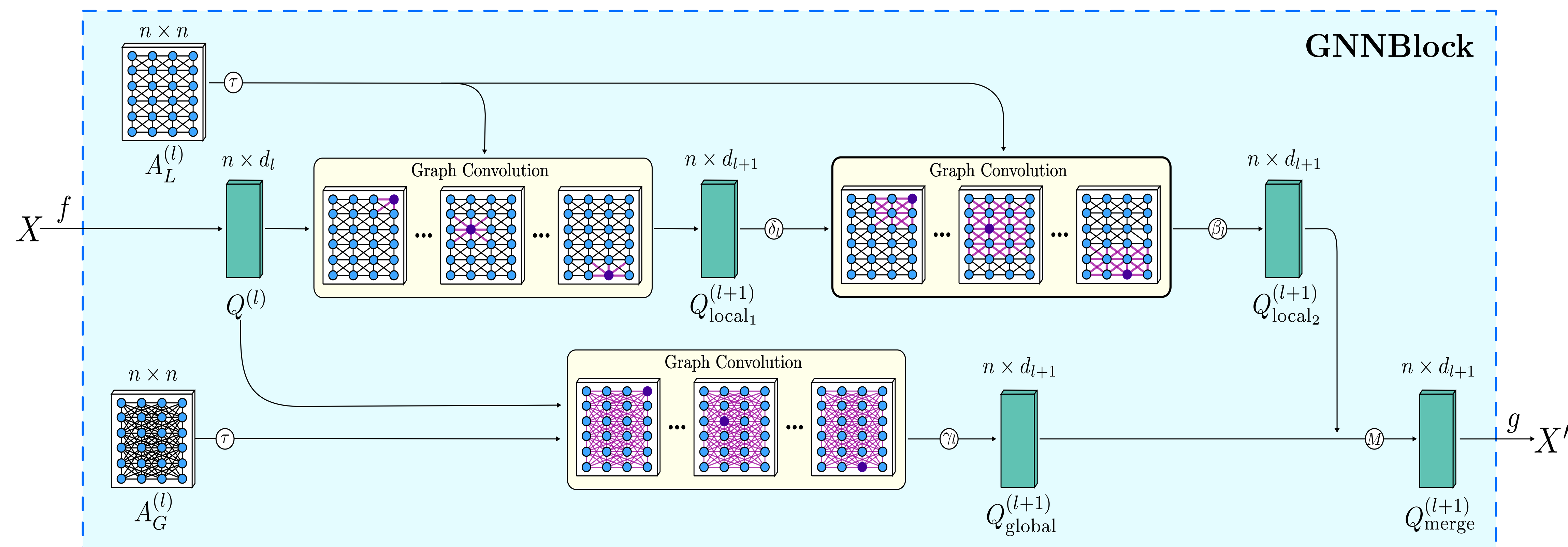


(a) DUNet (b) OCNet (c) HRNet+OCR

Table: Loss of spatial precision, generally displayed on the segmented objects' boundaries

Architecture for SS





$$\text{GNNBlock}(X) = g\Big(M\Big(\beta_l\big(\tau(A_L^{(l)})\delta_l\big(\tau(A_L^{(l)})f(X)Z^{(l)}_{\text{local}_1}\big)Z^{(l)}_{\text{local}_2}\big), \gamma_l\big(\tau(A_G^{(l)})f(X)Z^{(l)}_{\text{global}}\big)\Big)\Big) = X' \qquad (1)$$

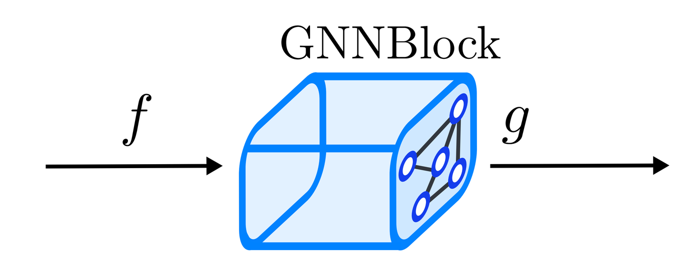
Methodology

Graph convolutional network, creating and deleting edges and updating features values.

$$H^{(l+1)} = \sigma_l\left(\tau\left(A^{(l)}\right)H^{(l)}W^{(l)}\right), \qquad (2)$$

$$\tau\left(A^{(l)}\right) = \left(\hat{D}^{(l)}\right)^{-\frac{1}{2}}\left(A^{(l)} + I_n\right)\left(\hat{D}^{(l)}\right)^{-\frac{1}{2}}, \qquad (3)$$

$$\hat{D}^{(l)} = D^{(l)} + I_n, \qquad (4)$$



$$X' = g(\sigma_l(\tau(A^{(l)})f(X)W^{(l)}))$$

f :transf. original space $\longrightarrow$ original space
g :transf. graph space $\longrightarrow$ graph space

Loss Functions

$$\mathcal{L}_{cross\_ss} = -\frac{1}{N}\sum_{i=1}^{N}\alpha_i \log P(s = s_i|X;\phi), \qquad (5)$$

$$\mathcal{L}_{iou\_ss} = 1 - \sum_{i=1}^{N}\frac{P_i \cap S_i}{P_i \cup S_i}. \qquad (6)$$

$$\mathcal{L}_{ss} = \psi_1\mathcal{L}_{cross\_ss} + \psi_2\mathcal{L}_{iou\_ss}. \qquad (7)$$

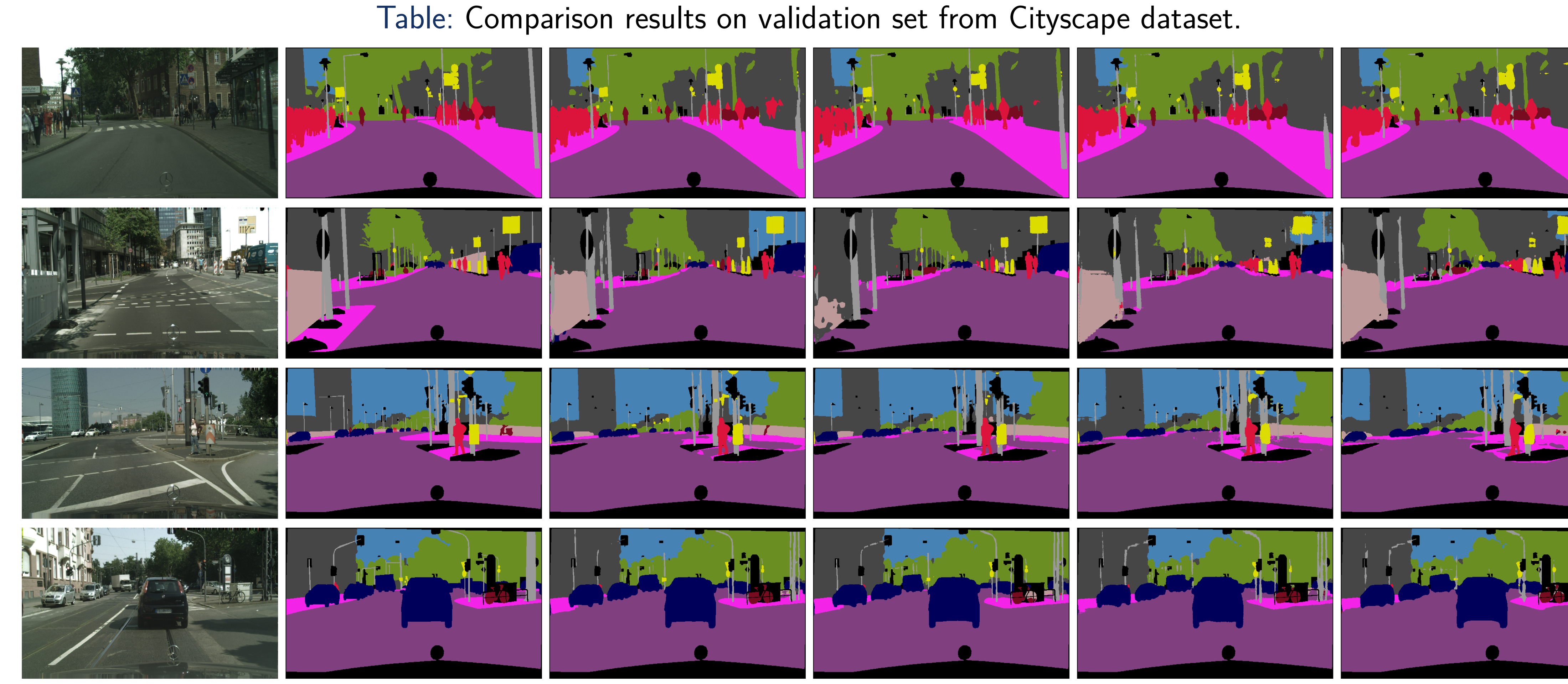
Quantitative Results

Table: IoU results on Cityscapes validation set for semantic segmentation, using 11 classes and with resize of 384×768.

Model	Sky	Building	Road	Sidewalk	Fence	Vegetat.	Pole	Car	Sign	Person	Cyclist	mIoU
ParseNet	90.76	85.20	92.01	63.49	39.00	88.21	46.82	89.20	61.90	63.91	62.73	71.20
ESPNet	91.79	86.36	95.73	71.84	48.52	88.44	49.06	87.29	54.60	61.83	57.51	72.09
FC-DenseNet67	92.19	86.77	96.60	75.40	41.55	88.07	52.92	87.09	63.89	60.48	52.92	72.54
BiSeNet	91.64	87.42	96.48	75.41	44.05	89.07	39.50	89.34	58.63	66.81	63.08	72.86
ENet	91.63	87.48	96.44	75.34	48.44	89.23	43.14	89.24	56.04	65.13	63.70	73.26
ICNet	92.03	88.44	96.61	77.22	42.59	89.46	48.27	90.74	58.71	66.18	66.28	74.23
DeepLab v3	92.82	89.02	96.74	78.13	41.00	90.81	49.74	91.02	64.48	66.52	66.98	75.21
PSPNet	91.94	89.93	96.94	78.37	53.64	90.19	43.47	92.12	64.40	70.71	70.94	76.61
DANet	92.25	90.26	97.25	79.95	51.33	90.60	45.20	92.50	66.38	71.47	71.25	77.13
AdapNet++	93.07	89.46	97.06	80.03	49.46	90.58	52.10	92.22	66.26	72.88	70.62	77.61
CCNet	90.97	89.01	96.59	77.36	42.49	91.36	58.24	91.22	71.18	74.80	70.57	77.62
OCNet	92.72	90.73	97.39	80.80	54.58	90.86	45.60	92.64	67.35	71.81	71.98	77.86
DUNet	93.33	91.05	97.28	80.18	55.15	91.35	53.70	92.89	68.33	73.33	72.29	78.99
GNNBlock* (Our)	**93.10**	**90.08**	**96.96**	**79.13**	**48.61**	**90.82**	**59.05**	**92.58**	**72.86**	**74.96**	**72.07**	**79.11**
HRNet	94.56	90.98	97.48	82.46	50.27	92.35	61.57	93.96	73.14	78.55	75.44	80.98

Qualitative Results

Table: Comparison results on validation set from Cityscape dataset.



(a) Image (b) Ground-truth (c) GNNNBlock* (d) HRNet (e) DUNet (f) OCNet