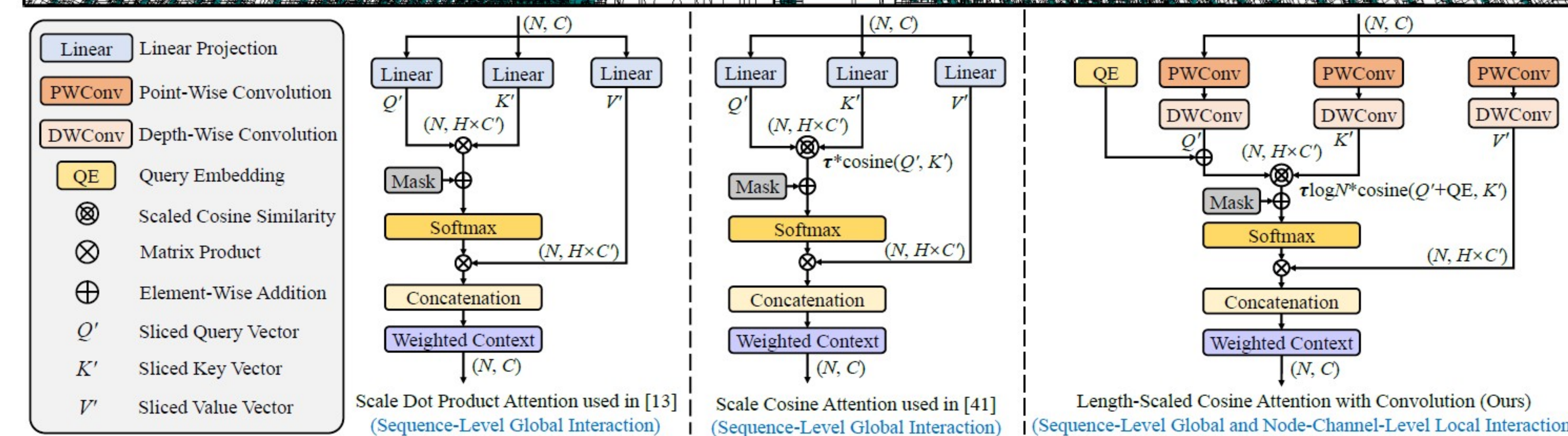
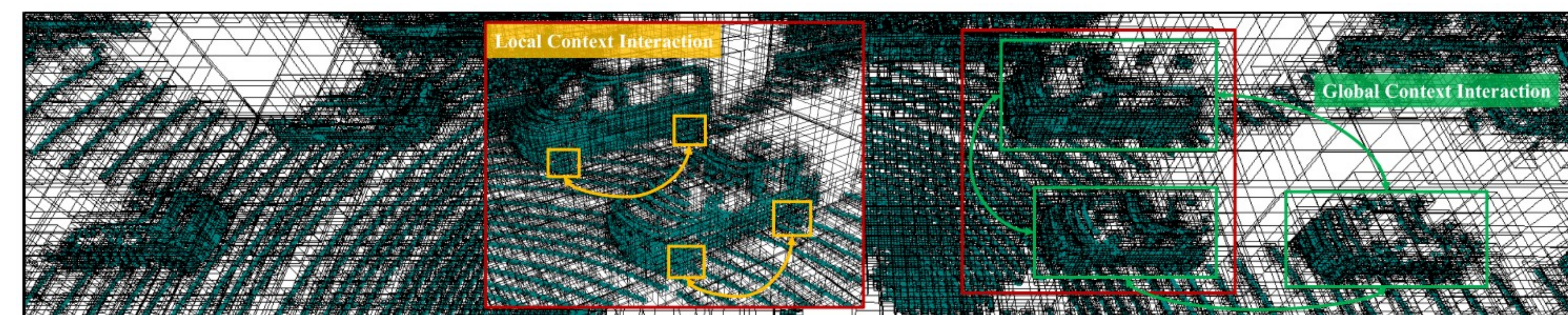


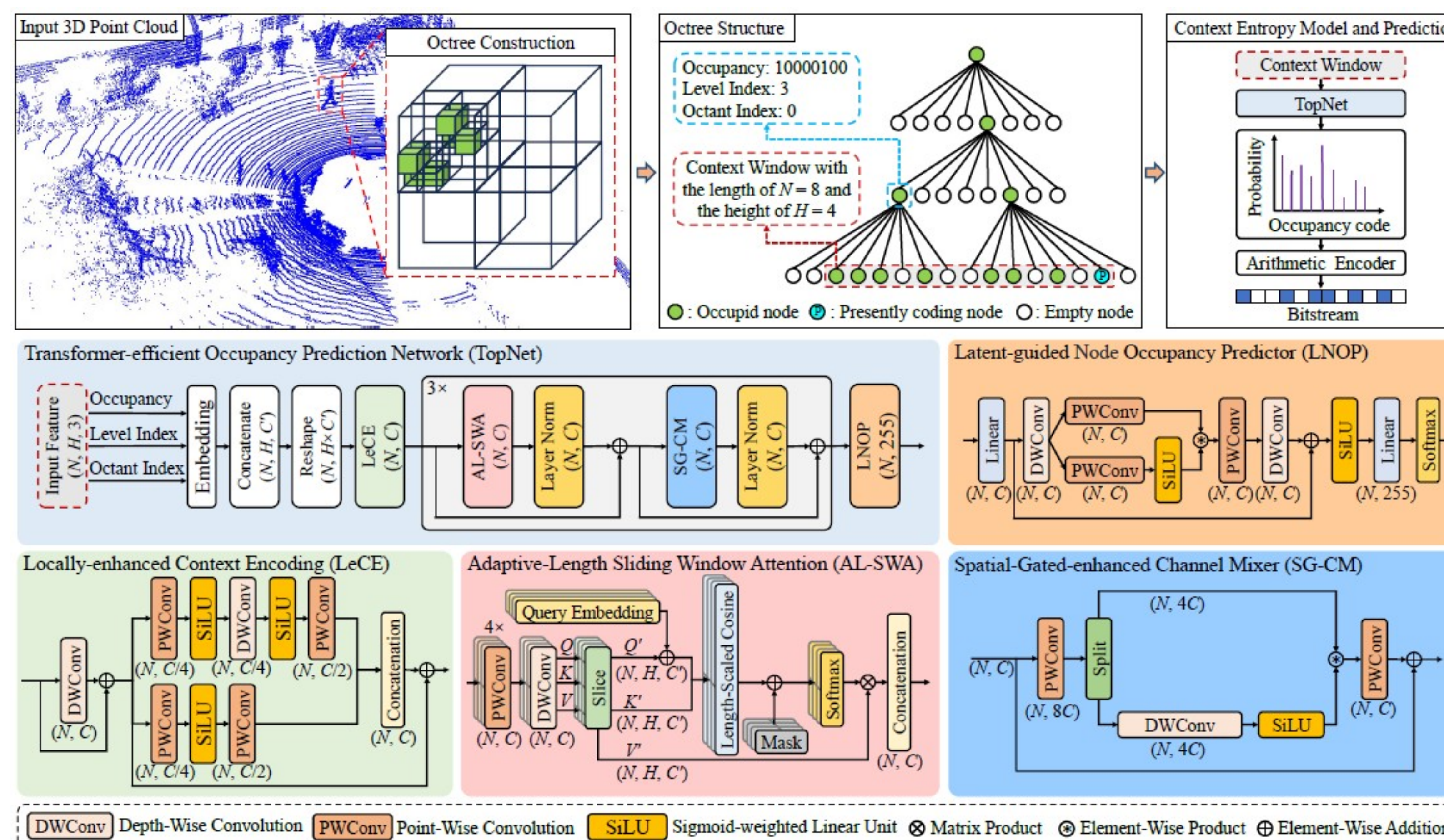
TopNet: Transformer-Efficient Occupancy Prediction Network for Octree-Structured Point Cloud Geometry Compression

Xinjie Wang*, Yifan Zhang, Ting Liu, Xinpu Liu, Ke Xu, Jianwei Wan, Yulan Guo, Hanyun Wang[†]

Motivation

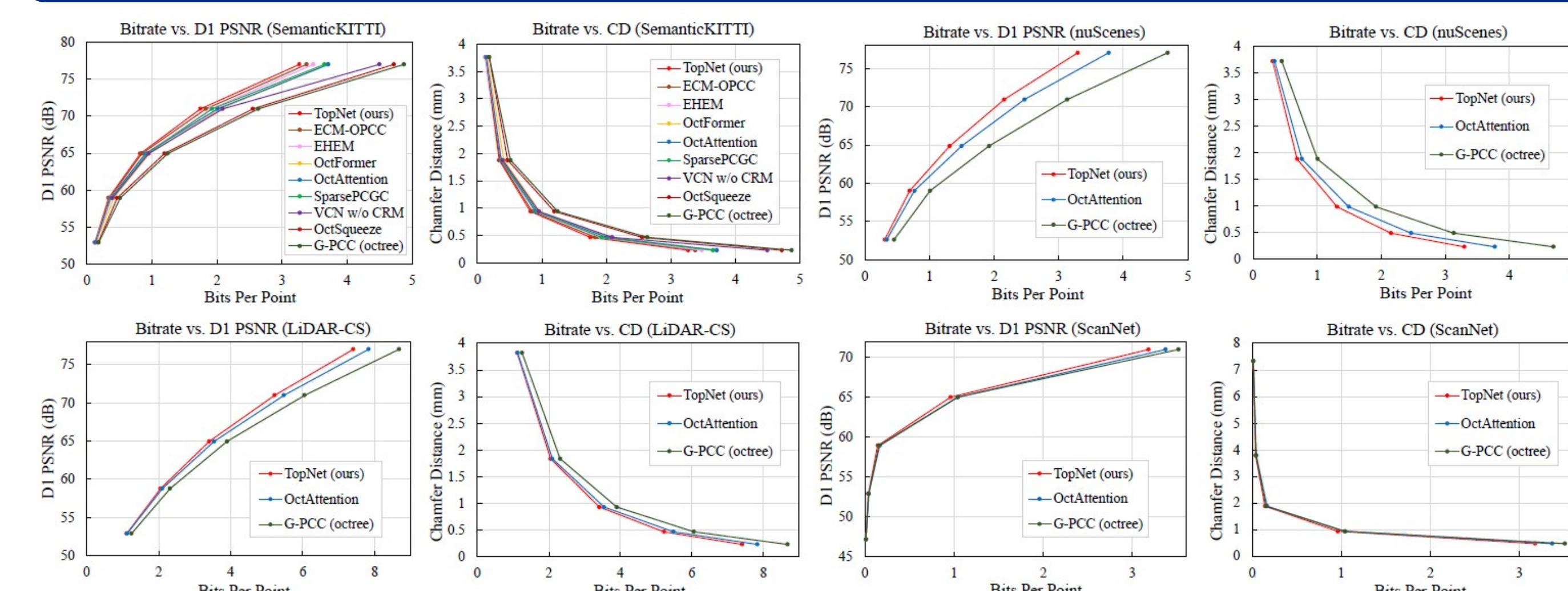


Methods

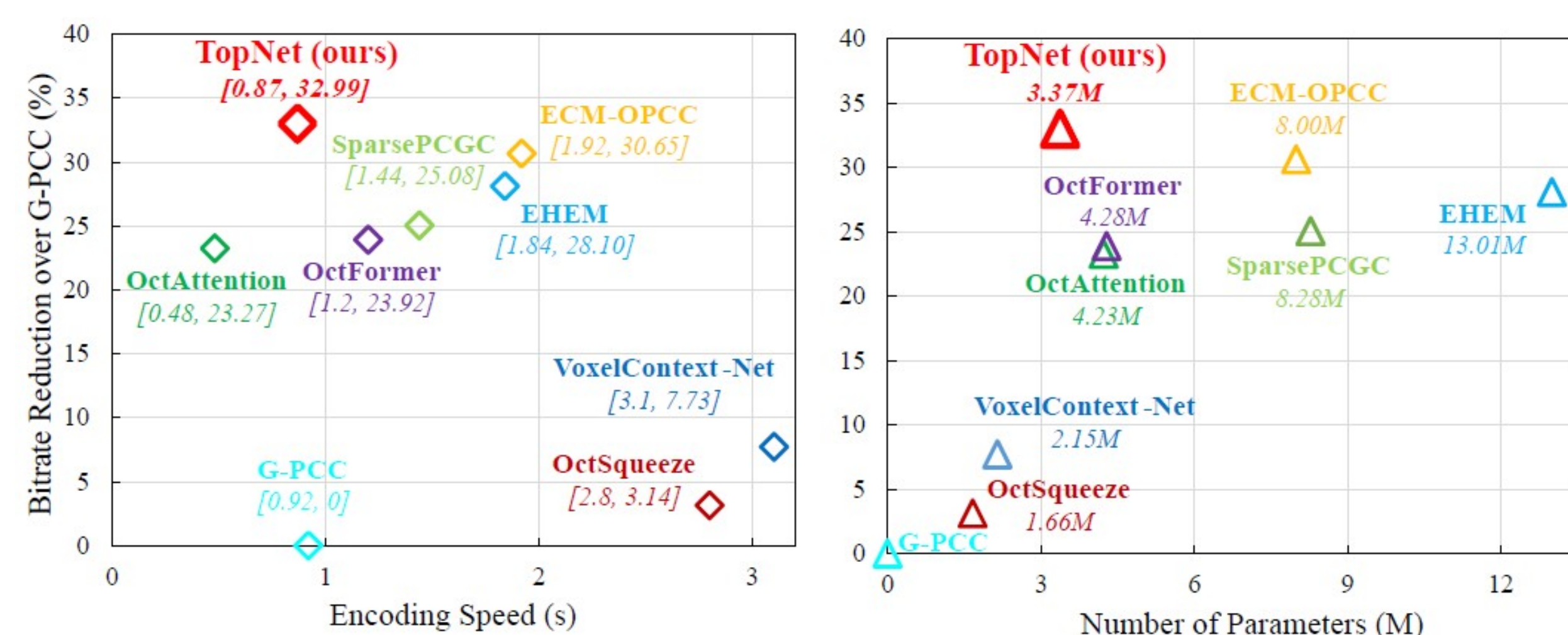


- LeCE for enhancing the **translation-invariance** of the octree nodes
- AL-SWA for capturing both **global** and **local** dependencies while **adaptively** adjusting attention weights based on input window length
- SG-CM for **efficient feature aggregation** from ancestors and siblings
- LNOP for improving prediction accuracy of spatially **adjacent** octree nodes

Qualitative & Quantitative Results

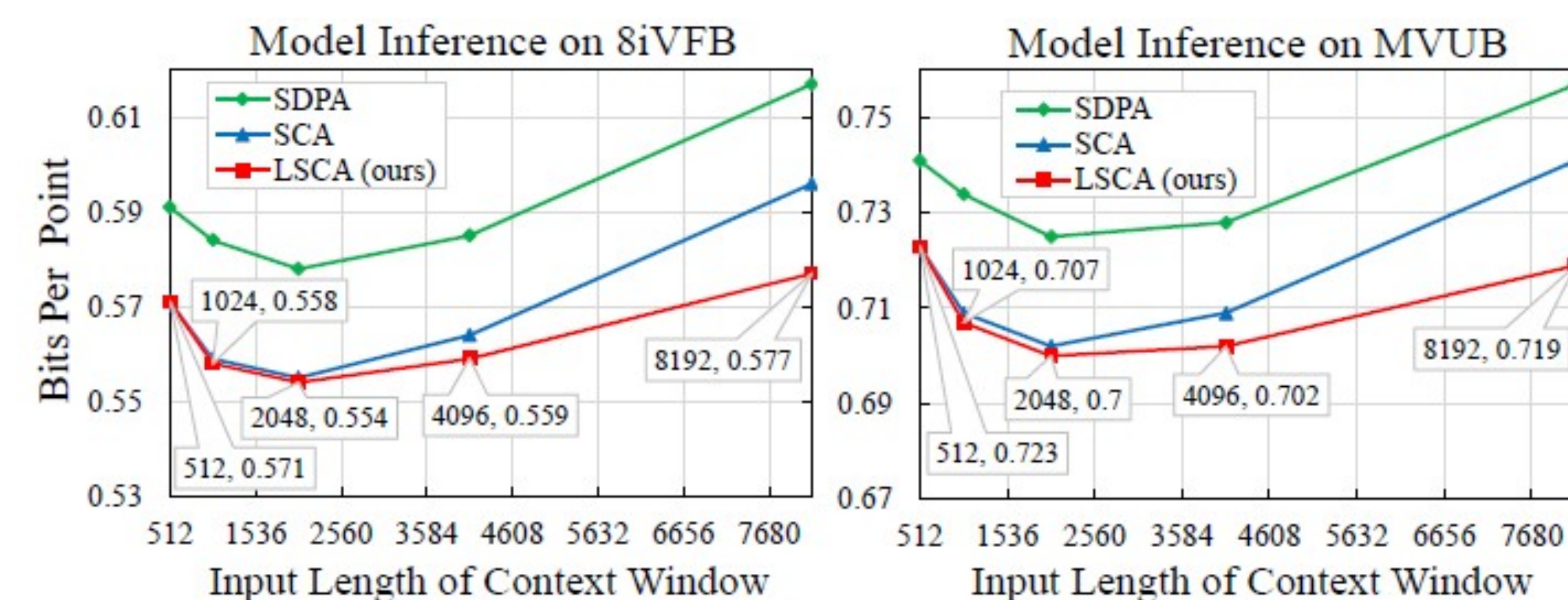


Point Cloud	G-PCC [32]	OctAttention [13]	VoxelDNN [26]	MSVoxelDNN [27]	ECM-OPCC [20]	TopNet (ours)
8i Voxelized Full Bodies (8iVFB) Dataset						
Boxer_vox9	0.96	0.60 (-37.50%)	0.76 (-20.83%)	0.70 (-25.53%)	0.53 (-44.79%)	0.51 (-46.88%)
Boxer_vox10	0.94	0.59 (-37.23%)	0.64 (-32.63%)	0.73 (-23.16%)	0.55 (-42.11%)	0.53 (-44.21%)
Loot_vox10	0.95	0.62 (-34.74%)	0.73 (-33.03%)	0.87 (-20.18%)	0.66 (-39.45%)	0.65 (-40.37%)
Redandblack_vox10	1.09	0.73 (-33.03%)	0.73 (-33.03%)	0.87 (-20.18%)	0.66 (-39.45%)	0.65 (-40.37%)
Thaidancer_vox9	0.99	0.64 (-35.35%)	0.81 (-18.18%)	-	0.59 (-40.40%)	0.56 (-43.43%)
Thaidancer_vox10	0.99	0.65 (-34.34%)	-	0.85 (-14.14%)	0.58 (-41.41%)	0.57 (-42.42%)
Avg.	0.99	0.64 (-35.35%)	0.73 (-26.26%)	0.79 (-20.20%)	0.58 (-41.41%)	0.56 (-43.43%)
Microsoft Voxelized Upper Bodies (MVUB) Dataset						
Phil_vox9	1.23	0.83 (-32.52%)	0.92 (-25.20%)	-	0.79 (-35.77%)	0.78 (-36.59%)
Phil_vox10	1.07	0.79 (-26.17%)	0.83 (-22.43%)	1.02 (-4.67%)	0.76 (-28.97%)	0.73 (-31.78%)
Ricardo_vox9	1.04	0.72 (-30.77%)	0.72 (-30.77%)	-	0.76 (-26.92%)	0.67 (-35.58%)
Ricardo_vox10	1.07	0.72 (-32.71%)	0.75 (-29.91%)	0.95 (-11.21%)	0.69 (-35.51%)	0.64 (-40.19%)
Avg.	1.10	0.76 (-30.91%)	0.81 (-26.36%)	0.99 (-10.00%)	0.75 (-31.82%)	0.71 (-35.45%)



Ablation Studies & Complexity Analysis

Model	LeCE	AL-SWA	SG-CM	LNOP	BPP (↓) on 8iVFB	BPP (↓) on MVUB	Param.
A					0.626	0.745	3.893M
B	✓				0.592 (-5.43%)	0.744 (-0.13%)	3.943M
C	✓	✓			0.565 (-9.74%)	0.716 (-3.89%)	3.953M
D	✓	✓	✓		0.562 (-10.22%)	0.712 (-4.43%)	3.169M
E	✓	✓	✓	✓	0.558 (-10.86%)	0.707 (-5.10%)	3.369M



Method	Enc.	Dec.	Train.	Param.	Mem.	Mod.
Lidar Point Cloud Datasets						
OctAttention [13]	0.19s	15.4s	48h	4.23M	0.37GB	28.0MB
TopNet (ours)	0.38s	14.9s	43h	3.37M	0.33GB	13.5MB
Human Body Point Cloud Datasets						
OctAttention [13]	0.29s	301s	9h	4.23M	0.42GB	28.0MB
TopNet (ours)	0.62s	292s	8h	3.37M	0.39GB	13.5MB