Few-shot Structure-Informed Machinery Part Segmentation with Foundation Models and Graph Neural Networks
Supplementary Material

Michael Schwingshackl Fabio F. Oberweger Markus Murschitz
AIT Austrian Institute of Technology
Center for Vision and Control
{michael.schwingshackl, fabio.oberweger, markus.murschitz}@ait.ac.at

A. Hyperparameters

Hyperparameter	Short	Description
NMS Radius	NR	SuperPoint non-maximum-supresion box radius for merging close points. Range: [2 - 6]
Min Points	I	Minumum amount of generated interest points for an image. Range: [512 - 2048]
Graph Neighbors	k	Amount of neighbor connections for each node. Range: [8 - 32]
Hidden Layers	HD	Depth of Hidden Layers in the Graph Classifier. Range: [256 - 1024]
Integration Layers	ID	Depth of Integration Layers in the Graph Classifier. Range: [128 - 512]
SP Threshold	SPT	SuperPoint threshold value for each interest point quality score. Range: [$1.0e^{-04}$ - $5.0e^{-04}$]
Model Type	MT	Graph convolution layer type. Values: [GAT, GAN, SAGE]
Dropout	DR	Dropout value for the Graph Classifier. Range: [0.1 - 0.3]
Dropout Edge	DRE	Edge Dropout value for the Graph Classifier. Range: 0.3 - 0.8]
SAM Prompt	SP	SAM prompt type for the Segmentation part. Values: [Point(P), Box(B), Point Box(PB)]
Point Threshold	PT	Threshold values for the point's Mahalanobis distance. Range: [0.6 - 1.0]
Box Threshold	BT	Threshold values for the bounding boxes's Mahalanobis distance. Range: [0.6 - 1.0]
SAM point samples	SPS	SAM input point amount for each class: Range [5 - 20]

Table 1. List of tunable hyperparameters, including description and ranges.

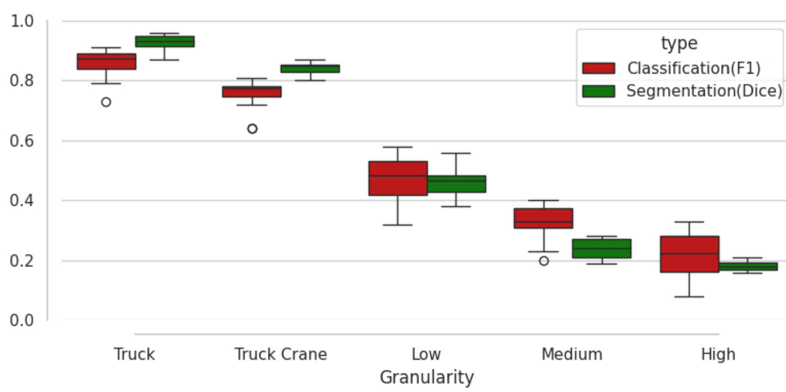

Figure 1. Boxplot visualization of all hyperparameter tuning runs, demonstrating the slight but noticeable differences between various parameter selections at each granularity level. The graph node classification F1 performance is represented in red, and the segmentation Dice score is displayed in green.

A.1. Hyperparameter Tuning Classification Part

Hyperparameters										Metrics
RUN	**NR**	***I***	***k***	**HD**	**ID**	**SPT**	**MT**	**DR**	**DRE**	**$F1_M$**
0	2	512	8	1024	128	$1.0e^{-04}$	GNN	0.2	0.5	0.85 ± 0.05
1	4	2048	8	512	256	$2.0e^{-04}$	SAGE	0.2	0.3	0.91 ± 0.04
2	2	1024	32	256	512	$2.0e^{-04}$	GAT	0.1	0.8	0.84 ± 0.06
3	4	512	32	512	128	$1.0e^{-04}$	SAGE	0.2	0.8	0.91 ± 0.04
4	2	512	8	1024	128	$1.0e^{-04}$	SAGE	0.1	0.8	0.89 ± 0.04
5	6	512	16	512	128	$1.0e^{-04}$	GAT	0.1	0.5	0.84 ± 0.06
6	4	1024	32	1024	128	$1.0e^{-04}$	GNN	0.2	0.8	0.81 ± 0.04
7	2	1024	16	1024	128	$2.0e^{-04}$	SAGE	0.3	0.5	0.89 ± 0.04
8	2	1024	8	512	128	$5.0e^{-04}$	GNN	0.2	0.8	0.87 ± 0.04
9	4	1024	16	1024	512	$5.0e^{-04}$	GAT	0.3	0.3	0.83 ± 0.07
10	2	2048	16	1024	128	$1.0e^{-04}$	SAGE	0.3	0.3	0.9 ± 0.04
11	6	2048	32	512	512	$1.0e^{-04}$	GNN	0.2	0.5	0.87 ± 0.05
12	2	1024	8	512	512	$2.0e^{-04}$	GAT	0.1	0.5	0.87 ± 0.05
13	4	1024	8	1024	512	$1.0e^{-04}$	GNN	0.3	0.5	0.88 ± 0.04
14	6	1024	32	256	128	$1.0e^{-04}$	GAT	0.2	0.8	0.73 ± 0.07
15	4	512	8	1024	256	$1.0e^{-04}$	SAGE	0.1	0.8	0.89 ± 0.05
16	6	2048	8	512	128	$5.0e^{-04}$	GAT	0.3	0.3	0.88 ± 0.05
17	6	2048	16	1024	256	$2.0e^{-04}$	GNN	0.3	0.3	0.88 ± 0.04
18	2	512	32	512	512	$5.0e^{-04}$	GAT	0.3	0.3	0.79 ± 0.06
19	2	1024	32	256	256	$1.0e^{-04}$	SAGE	0.1	0.8	0.9 ± 0.04

Table 2. 20 runs of graph classification hyperparameter tuning with ***TRUCK*** granularity. Best and worst runs are highlighted in green and red. Metric evaluated on 250 test samples(mean/std).

Hyperparameters										Metrics
RUN	**NR**	***I***	***k***	**HD**	**ID**	**SPT**	**MT**	**DR**	**DRE**	**$F1_M$**
0	6	512	16	512	256	$5.0e^{-04}$	GNN	0.2	0.8	0.74 ± 0.12
1	6	1024	16	512	512	$2.0e^{-04}$	GNN	0.2	0.3	0.77 ± 0.12
2	6	512	32	512	128	$5.0e^{-04}$	SAGE	0.1	0.3	0.81 ± 0.11
3	2	512	32	512	128	$1.0e^{-04}$	SAGE	0.1	0.8	0.8 ± 0.11
4	4	2048	32	256	128	$2.0e^{-04}$	GNN	0.2	0.5	0.76 ± 0.12
5	2	1024	8	256	256	$1.0e^{-04}$	GAT	0.1	0.5	0.78 ± 0.09
6	6	2048	8	512	512	$5.0e^{-04}$	GAT	0.3	0.8	0.78 ± 0.1
7	2	2048	32	256	512	$5.0e^{-04}$	GNN	0.2	0.5	0.76 ± 0.11
8	2	512	32	1024	128	$2.0e^{-04}$	GAT	0.3	0.8	0.64 ± 0.1
9	2	1024	32	1024	512	$2.0e^{-04}$	SAGE	0.2	0.3	0.81 ± 0.11
10	2	1024	8	256	512	$2.0e^{-04}$	SAGE	0.1	0.8	0.78 ± 0.1
11	2	2048	8	1024	512	$1.0e^{-04}$	SAGE	0.1	0.5	0.79 ± 0.1
12	4	1024	8	256	128	$5.0e^{-04}$	GNN	0.2	0.3	0.78 ± 0.11
13	6	512	16	1024	256	$1.0e^{-04}$	GNN	0.1	0.8	0.72 ± 0.11
14	4	512	32	256	256	$5.0e^{-04}$	SAGE	0.1	0.5	0.73 ± 0.12
15	4	1024	8	512	128	$2.0e^{-04}$	GNN	0.1	0.5	0.78 ± 0.1
16	4	1024	8	512	512	$5.0e^{-04}$	GAT	0.1	0.5	0.75 ± 0.1
17	6	512	16	512	256	$2.0e^{-04}$	GAT	0.3	0.5	0.64 ± 0.12
18	6	512	8	256	512	$5.0e^{-04}$	GNN	0.3	0.5	0.77 ± 0.12
19	2	1024	16	1024	128	$5.0e^{-04}$	SAGE	0.1	0.8	0.8 ± 0.1

Table 3. 20 runs of graph classification hyperparameter tuning with ***TRUCK CRANE*** granularity. Best and worst runs are highlighted in green and red. Metric evaluated on 250 test samples(mean/std).

Hyperparameters										Metrics
RUN	**NR**	***I***	***k***	**HD**	**ID**	**SPT**	**MT**	**DR**	**DRE**	$\mathbf{F1_M}$
0	4	1024	32	1024	256	$1.0e^{-04}$	GAT	0.1	0.3	0.32 ± 0.09
1	6	2048	8	256	128	$5.0e^{-04}$	GNN	0.2	0.5	0.43 ± 0.12
2	2	1024	8	1024	256	$5.0e^{-04}$	GNN	0.3	0.8	0.53 ± 0.15
3	4	1024	32	512	256	$5.0e^{-04}$	SAGE	0.2	0.3	0.58 ± 0.16
4	2	2048	16	256	256	$5.0e^{-04}$	SAGE	0.3	0.8	0.53 ± 0.16
5	6	1024	16	1024	256	$5.0e^{-04}$	GNN	0.1	0.5	0.42 ± 0.12
6	2	2048	16	256	256	$2.0e^{-04}$	GAT	0.1	0.8	0.5 ± 0.14
7	4	512	16	256	128	$2.0e^{-04}$	GAT	0.3	0.5	0.42 ± 0.12
8	6	1024	32	512	128	$1.0e^{-04}$	GNN	0.2	0.3	0.44 ± 0.12
9	4	2048	8	512	256	$2.0e^{-04}$	SAGE	0.3	0.3	0.54 ± 0.16
10	4	512	32	256	256	$5.0e^{-04}$	GNN	0.1	0.8	0.37 ± 0.11
11	4	512	8	256	512	$1.0e^{-04}$	GAT	0.2	0.8	0.36 ± 0.11
12	2	1024	16	512	512	$2.0e^{-04}$	GNN	0.1	0.8	0.53 ± 0.15
13	6	512	8	1024	512	$5.0e^{-04}$	GAT	0.2	0.3	0.35 ± 0.11
14	6	512	8	512	256	$1.0e^{-04}$	GNN	0.3	0.8	0.48 ± 0.13
15	6	1024	16	512	256	$1.0e^{-04}$	GNN	0.3	0.5	0.49 ± 0.13
16	2	2048	8	256	128	$1.0e^{-04}$	GNN	0.1	0.8	0.53 ± 0.15
17	2	512	32	256	128	$2.0e^{-04}$	GAT	0.2	0.5	0.41 ± 0.12
18	4	2048	8	1024	256	$2.0e^{-04}$	SAGE	0.1	0.5	0.53 ± 0.15
19	2	512	8	512	256	$1.0e^{-04}$	SAGE	0.3	0.5	0.55 ± 0.15

Table 4. 20 runs of graph classification hyperparameter tuning with ***LOW*** granularity. Best and worst runs are highlighted in green and red. Metric evaluated on 250 test samples(mean/std).

Hyperparameters										Metrics
RUN	**NR**	***I***	***k***	**HD**	**ID**	**SPT**	**MT**	**DR**	**DRE**	$\mathbf{F1_M}$
0	6	2048	8	512	512	$1.0e^{-04}$	GAT	0.1	0.5	0.35 ± 0.1
1	4	2048	32	1024	512	$1.0e^{-04}$	GNN	0.3	0.3	0.34 ± 0.12
2	6	512	8	1024	256	$1.0e^{-04}$	SAGE	0.3	0.5	0.4 ± 0.14
3	6	1024	32	256	128	$2.0e^{-04}$	GNN	0.1	0.8	0.23 ± 0.08
4	2	1024	16	512	512	$1.0e^{-04}$	GAT	0.2	0.3	0.31 ± 0.1
5	2	512	16	256	256	$5.0e^{-04}$	SAGE	0.2	0.5	0.39 ± 0.14
6	2	2048	8	256	256	$1.0e^{-04}$	GNN	0.3	0.5	0.37 ± 0.14
7	4	1024	8	512	512	$1.0e^{-04}$	GNN	0.1	0.8	0.33 ± 0.1
8	2	1024	32	1024	256	$2.0e^{-04}$	SAGE	0.3	0.5	0.4 ± 0.15
9	2	512	16	512	512	$1.0e^{-04}$	SAGE	0.1	0.8	0.32 ± 0.1
10	4	1024	8	256	512	$1.0e^{-04}$	SAGE	0.2	0.3	0.37 ± 0.13
11	6	512	16	1024	128	$2.0e^{-04}$	SAGE	0.3	0.3	0.33 ± 0.11
12	6	1024	8	512	128	$1.0e^{-04}$	SAGE	0.3	0.8	0.39 ± 0.14
13	4	1024	16	1024	128	$2.0e^{-04}$	GNN	0.1	0.3	0.31 ± 0.1
14	6	1024	16	256	512	$5.0e^{-04}$	GNN	0.3	0.3	0.31 ± 0.11
15	2	512	8	512	512	$5.0e^{-04}$	GNN	0.1	0.5	0.29 ± 0.1
16	4	512	8	256	128	$2.0e^{-04}$	GNN	0.2	0.5	0.31 ± 0.1
17	6	1024	16	256	512	$5.0e^{-04}$	SAGE	0.2	0.8	0.28 ± 0.08
18	6	1024	16	1024	256	$5.0e^{-04}$	GAT	0.2	0.8	0.2 ± 0.06
19	4	512	8	1024	512	$1.0e^{-04}$	SAGE	0.3	0.8	0.39 ± 0.13

Table 5. 20 runs of graph classification hyperparameter tuning with ***MEDIUM*** granularity. Best and worst runs are highlighted in green and red. Metric evaluated on 250 test samples(mean/std).

Hyperparameters										Metrics
RUN	**NR**	***I***	***k***	**HD**	**ID**	**SPT**	**MT**	**DR**	**DRE**	$\mathbf{F1_M}$
0	4	512	16	512	512	$1.0e^{-04}$	SAGE	0.2	0.5	0.32 ± 0.13
1	2	2048	16	256	128	$1.0e^{-04}$	SAGE	0.1	0.8	0.29 ± 0.11
2	6	512	32	512	256	$1.0e^{-04}$	GAT	0.3	0.8	0.1 ± 0.03
3	4	2048	8	512	512	$5.0e^{-04}$	GAT	0.3	0.3	0.27 ± 0.09
4	6	512	16	256	512	$2.0e^{-04}$	GAT	0.1	0.8	0.14 ± 0.04
5	6	512	8	1024	256	$5.0e^{-04}$	GAT	0.1	0.3	0.13 ± 0.04
6	2	2048	8	512	128	$2.0e^{-04}$	GNN	0.3	0.5	0.28 ± 0.1
7	4	512	8	256	512	$1.0e^{-04}$	GNN	0.1	0.3	0.23 ± 0.07
8	4	512	32	256	128	$2.0e^{-04}$	GAT	0.2	0.3	0.19 ± 0.06
9	6	1024	16	512	512	$5.0e^{-04}$	GAT	0.2	0.5	0.22 ± 0.07
10	2	512	32	1024	512	$1.0e^{-04}$	GAT	0.1	0.3	0.08 ± 0.03
11	4	512	16	1024	128	$1.0e^{-04}$	GNN	0.3	0.8	0.17 ± 0.06
12	2	512	16	256	256	$5.0e^{-04}$	GAT	0.3	0.5	0.22 ± 0.08
13	2	2048	8	1024	512	$2.0e^{-04}$	GNN	0.2	0.8	0.29 ± 0.12
14	4	512	16	256	256	$1.0e^{-04}$	SAGE	0.1	0.5	0.33 ± 0.12
15	6	512	16	1024	512	$5.0e^{-04}$	SAGE	0.3	0.5	0.22 ± 0.07
16	4	1024	8	256	128	$2.0e^{-04}$	GAT	0.3	0.5	0.24 ± 0.08
17	4	1024	8	1024	256	$1.0e^{-04}$	SAGE	0.1	0.8	0.3 ± 0.12
18	6	1024	32	512	128	$1.0e^{-04}$	GAT	0.1	0.8	0.13 ± 0.04
19	2	2048	8	512	256	$2.0e^{-04}$	GAT	0.3	0.5	0.27 ± 0.09

Table 6. 20 runs of graph classification hyperparameter tuning with ***HIGH*** granularity. Best and worst runs are highlighted in green and red. Metric evaluated on 250 test samples(mean/std).

A.2. Hyperparameter Tuning Segmentation Part

Hyperparameters					Metrics
RUN	**SP**	**BT**	**PT**	**SPS**	**DICE**
0	P	1.0	0.6	15	0.92 ± 0.06
1	PB	0.8	0.8	10	0.94 ± 0.05
2	B	0.6	0.6	15	0.87 ± 0.09
3	PB	1.0	1.0	20	0.96 ± 0.03
4	P	1.0	0.8	15	0.93 ± 0.07
5	PB	0.6	0.6	10	0.9 ± 0.07
6	B	1.0	1.0	20	0.96 ± 0.03
7	B	0.6	1.0	20	0.87 ± 0.09
8	B	1.0	1.0	10	0.95 ± 0.08
9	P	1.0	1.0	20	0.95 ± 0.03
10	PB	1.0	1.0	20	0.96 ± 0.03
11	B	0.6	1.0	20	0.87 ± 0.1
12	P	1.0	1.0	20	0.94 ± 0.06
13	PB	0.6	0.6	10	0.89 ± 0.07
14	P	1.0	0.6	15	0.92 ± 0.04
15	P	1.0	0.6	15	0.92 ± 0.05
16	P	0.6	1.0	15	0.95 ± 0.03
17	PB	0.8	0.6	10	0.94 ± 0.04
18	B	0.8	1.0	20	0.93 ± 0.07
19	P	0.8	0.6	10	0.92 ± 0.06

Table 7. 20 runs of segmentation hyperparameter tuning with ***TRUCK*** granularity. Best and worst runs are highlighted in green and red. Metric evaluated on 250 test samples(mean/std).

Hyperparameters					Metrics
RUN	**SP**	**BT**	**PT**	**SPS**	**DICE**
0	PB	0.8	0.6	20	0.86 ± 0.11
1	B	0.8	0.6	5	0.85 ± 0.13
2	PB	1.0	0.6	10	0.86 ± 0.11
3	PB	1.0	0.6	5	0.86 ± 0.12
4	PB	1.0	1.0	15	0.85 ± 0.14
5	P	0.6	0.8	20	0.85 ± 0.14
6	PB	0.6	0.6	15	0.84 ± 0.12
7	PB	1.0	0.8	5	0.86 ± 0.11
8	P	0.8	0.8	5	0.8 ± 0.16
9	B	0.8	1.0	5	0.83 ± 0.14
10	P	0.8	1.0	15	0.8 ± 0.17
11	B	0.6	0.6	5	0.81 ± 0.14
12	PB	0.8	0.6	10	0.87 ± 0.11
13	P	0.8	0.8	20	0.83 ± 0.15
14	PB	0.6	0.8	5	0.83 ± 0.12
15	PB	0.6	0.8	15	0.84 ± 0.12
16	PB	1.0	0.8	15	0.85 ± 0.13
17	P	1.0	0.8	10	0.85 ± 0.14
18	PB	1.0	0.8	20	0.85 ± 0.12
19	PB	1.0	0.6	20	0.85 ± 0.13

Table 8. 20 runs of segmentation hyperparameter tuning with ***TRUCK CRANE*** granularity. Best and worst runs are highlighted in green and red. Metric evaluated on 250 test samples(mean/std).

Hyperparameters					Metrics
RUN	**SP**	**BT**	**PT**	**SPS**	**DICE**
0	PB	0.8	1.0	20	0.51 ± 0.14
1	PB	0.8	0.6	10	0.52 ± 0.14
2	P	0.6	0.6	15	0.46 ± 0.14
3	PB	0.6	0.6	10	0.56 ± 0.15
4	P	0.6	0.8	20	0.47 ± 0.14
5	P	1.0	0.6	20	0.49 ± 0.14
6	B	1.0	1.0	5	0.43 ± 0.13
7	P	0.8	0.6	5	0.4 ± 0.13
8	P	0.6	1.0	5	0.38 ± 0.12
9	P	0.8	0.8	10	0.44 ± 0.14
10	P	0.6	0.6	15	0.47 ± 0.14
11	P	1.0	1.0	10	0.41 ± 0.13
12	P	0.8	0.8	15	0.45 ± 0.14
13	P	1.0	0.8	20	0.47 ± 0.15
14	PB	1.0	1.0	15	0.48 ± 0.13
15	B	0.8	1.0	5	0.49 ± 0.15
16	P	0.8	0.8	10	0.42 ± 0.12
17	PB	1.0	1.0	20	0.47 ± 0.13
18	B	1.0	0.8	5	0.43 ± 0.14
19	B	1.0	0.8	20	0.43 ± 0.14

Table 9. 20 runs of segmentation hyperparameter tuning with ***LOW*** granularity. Best and worst runs are highlighted in green and red. Metric evaluated on 250 test samples(mean/std).

Hyperparameters					Metrics
RUN	**SP**	**BT**	**PT**	**SPS**	**DICE**
0	P	0.6	0.8	5	0.2 ± 0.07
1	P	0.8	1.0	10	0.21 ± 0.07
2	P	1.0	0.6	5	0.2 ± 0.08
3	P	0.8	0.8	10	0.21 ± 0.08
4	PB	0.8	0.8	10	0.28 ± 0.09
5	B	1.0	0.6	20	0.22 ± 0.09
6	PB	0.8	0.8	10	0.27 ± 0.09
7	P	0.6	0.6	5	0.19 ± 0.08
8	PB	0.8	1.0	10	0.27 ± 0.09
9	PB	0.8	0.6	15	0.27 ± 0.1
10	B	1.0	1.0	5	0.23 ± 0.07
11	B	0.6	0.8	5	0.23 ± 0.09
12	P	0.6	0.6	20	0.24 ± 0.08
13	PB	0.6	0.6	5	0.25 ± 0.11
14	P	1.0	0.6	10	0.21 ± 0.08
15	PB	0.6	0.8	20	0.27 ± 0.1
16	PB	1.0	0.8	10	0.27 ± 0.09
17	B	1.0	1.0	10	0.24 ± 0.08
18	PB	0.8	0.6	20	0.26 ± 0.1
19	P	0.6	0.8	20	0.24 ± 0.08

Table 10. 20 runs of segmentation hyperparameter tuning with ***MEDIUM*** granularity. Best and worst runs are highlighted in green and red. Metric evaluated on 250 test samples(mean/std).

Hyperparameters					Metrics
RUN	**SP**	**BT**	**PT**	**SPS**	**DICE**
0	B	0.6	0.6	15	0.18 ± 0.06
1	PB	1.0	0.6	10	0.2 ± 0.07
2	P	0.6	0.8	5	0.16 ± 0.05
3	B	1.0	0.6	20	0.17 ± 0.07
4	B	0.6	0.8	10	0.18 ± 0.06
5	P	1.0	0.8	15	0.17 ± 0.06
6	PB	0.6	0.6	15	0.21 ± 0.07
7	P	0.8	1.0	20	0.18 ± 0.06
8	B	0.8	1.0	20	0.18 ± 0.06
9	PB	1.0	0.6	20	0.19 ± 0.07
10	B	0.8	1.0	15	0.18 ± 0.06
11	PB	0.6	0.8	5	0.2 ± 0.06
12	PB	0.6	0.8	15	0.21 ± 0.07
13	B	1.0	1.0	5	0.17 ± 0.06
14	PB	1.0	1.0	5	0.19 ± 0.07
15	P	0.6	0.8	20	0.18 ± 0.07
16	P	0.8	0.8	20	0.18 ± 0.07
17	B	1.0	0.6	5	0.17 ± 0.06
18	P	0.8	0.8	10	0.16 ± 0.06
19	PB	0.8	0.8	15	0.21 ± 0.07

Table 11. 20 runs of segmentation hyperparameter tuning with ***HIGH*** granularity. Best and worst runs are highlighted in green and red. Metric evaluated on 250 test samples(mean/std).

B. Samples: *Truck*

Figure 2. ***Truck*** granularity graph node classification results for different train sample sizes. Each row shows the best five classification results for the given train sample size measured with *F1* score.

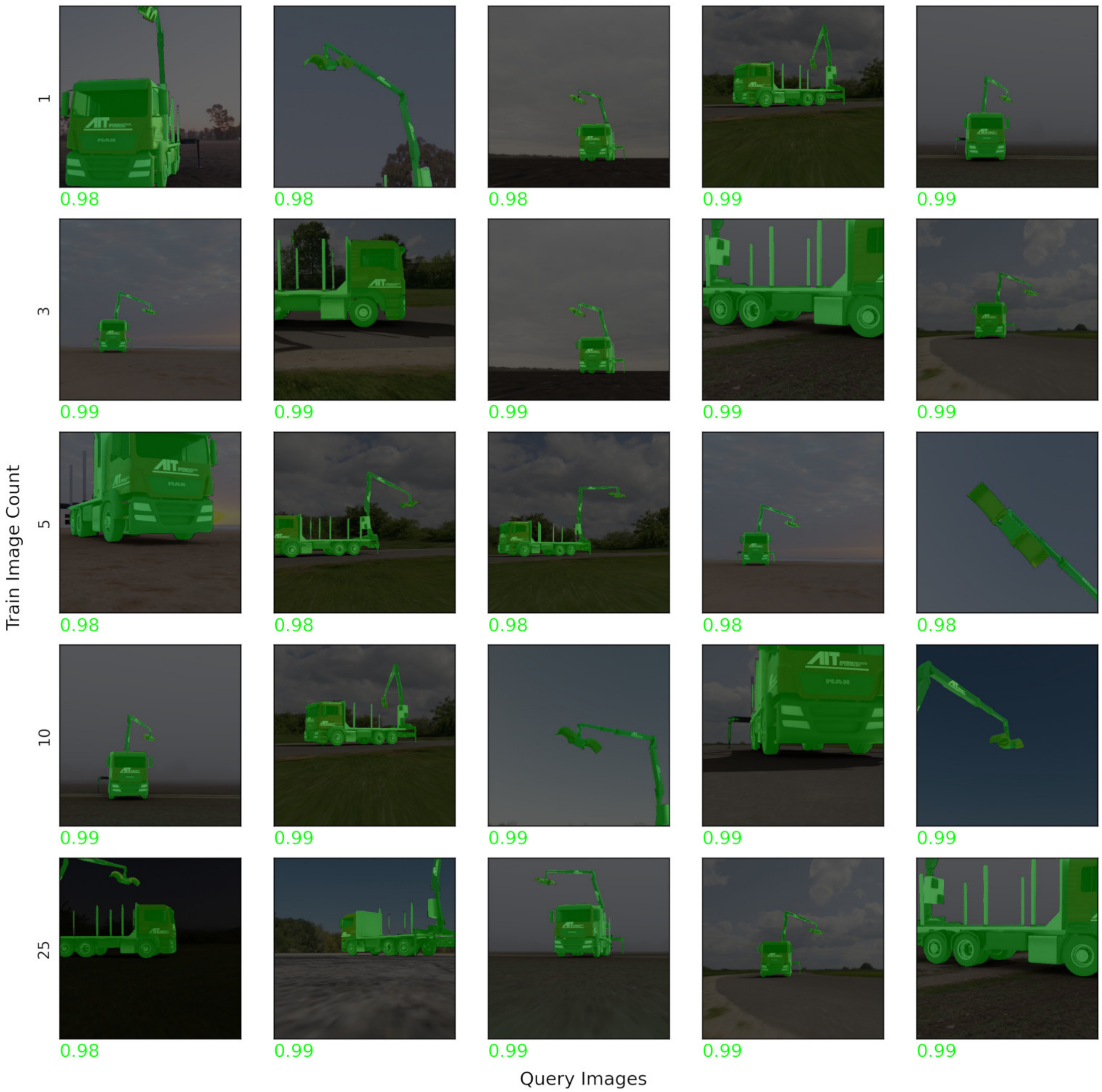

Figure 3. ***Truck*** granularity semantic segmentation results for different train sample sizes. Each row shows the best five segmentation results for the given train sample size measured with *dice* score.

C. Samples: *Truck Crane*

Figure 4. ***Truck Crane*** granularity graph node classification results for different train sample sizes. Each row shows the best five classification results for the given train sample size measured with *F1* score.

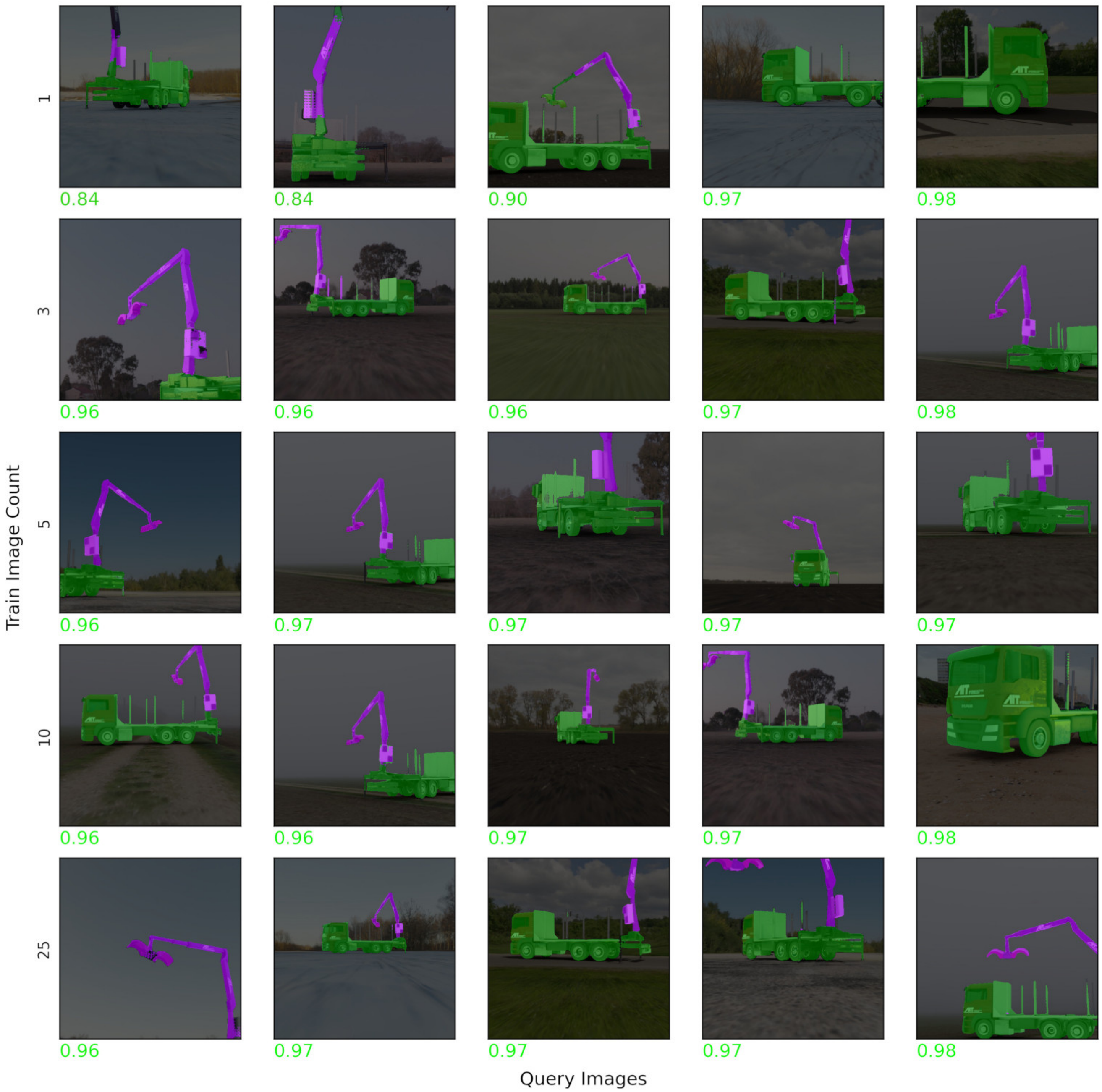

Figure 5. ***Truck Crane*** granularity semantic segmentation results for different train sample sizes. Each row shows the best five segmentation results for the given train sample size measured with *dice* score.

D. Samples: *Low*

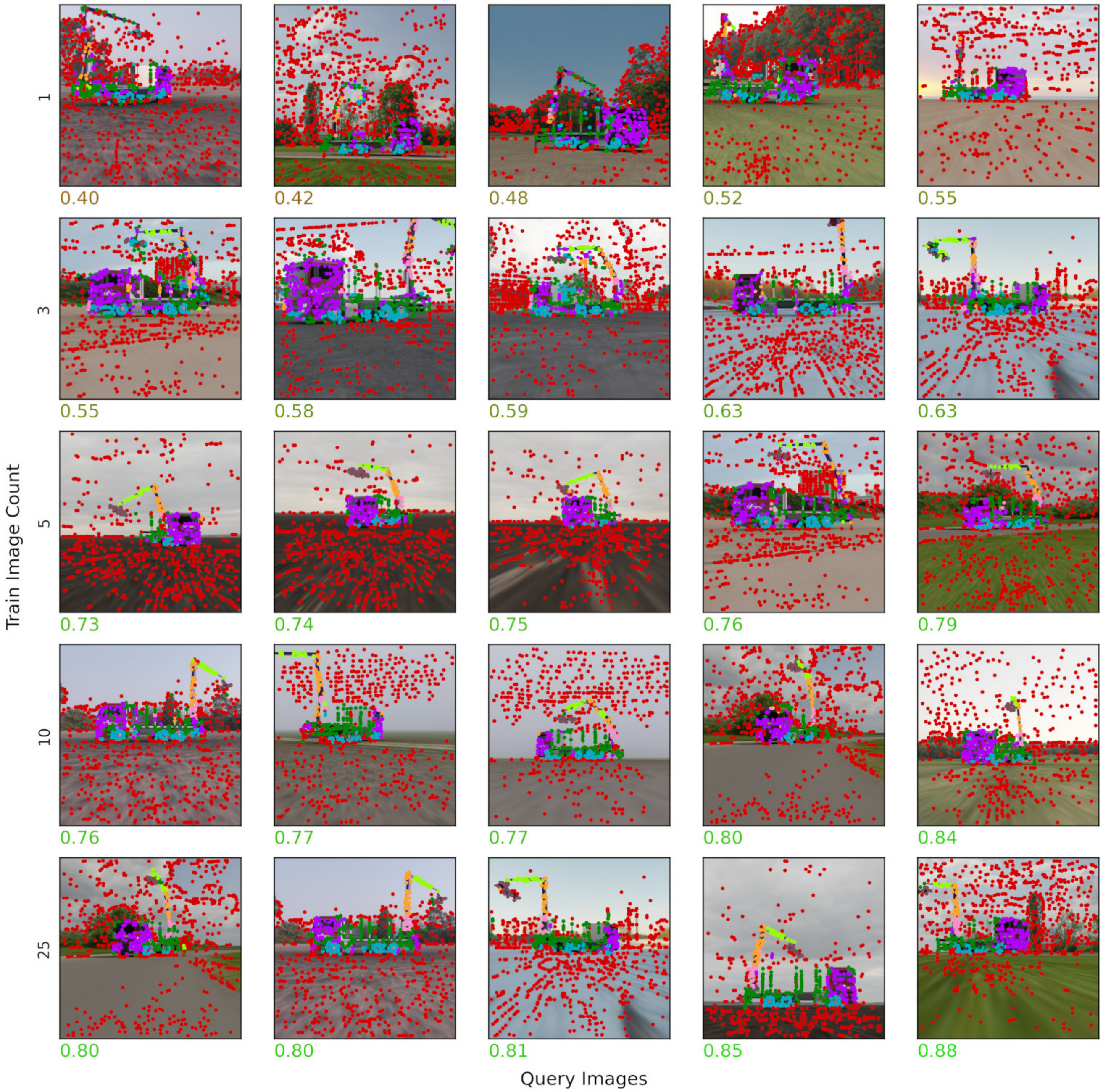

Figure 6. ***Low*** granularity graph node classification results for different train sample sizes. Each row shows the best five classification results for the given train sample size measured with *F1* score.

Figure 7. ***Low*** granularity semantic segmentation results for different train sample sizes. Each row shows the best five segmentation results for the given train sample size measured with *dice* score.

E. Samples: *Medium*

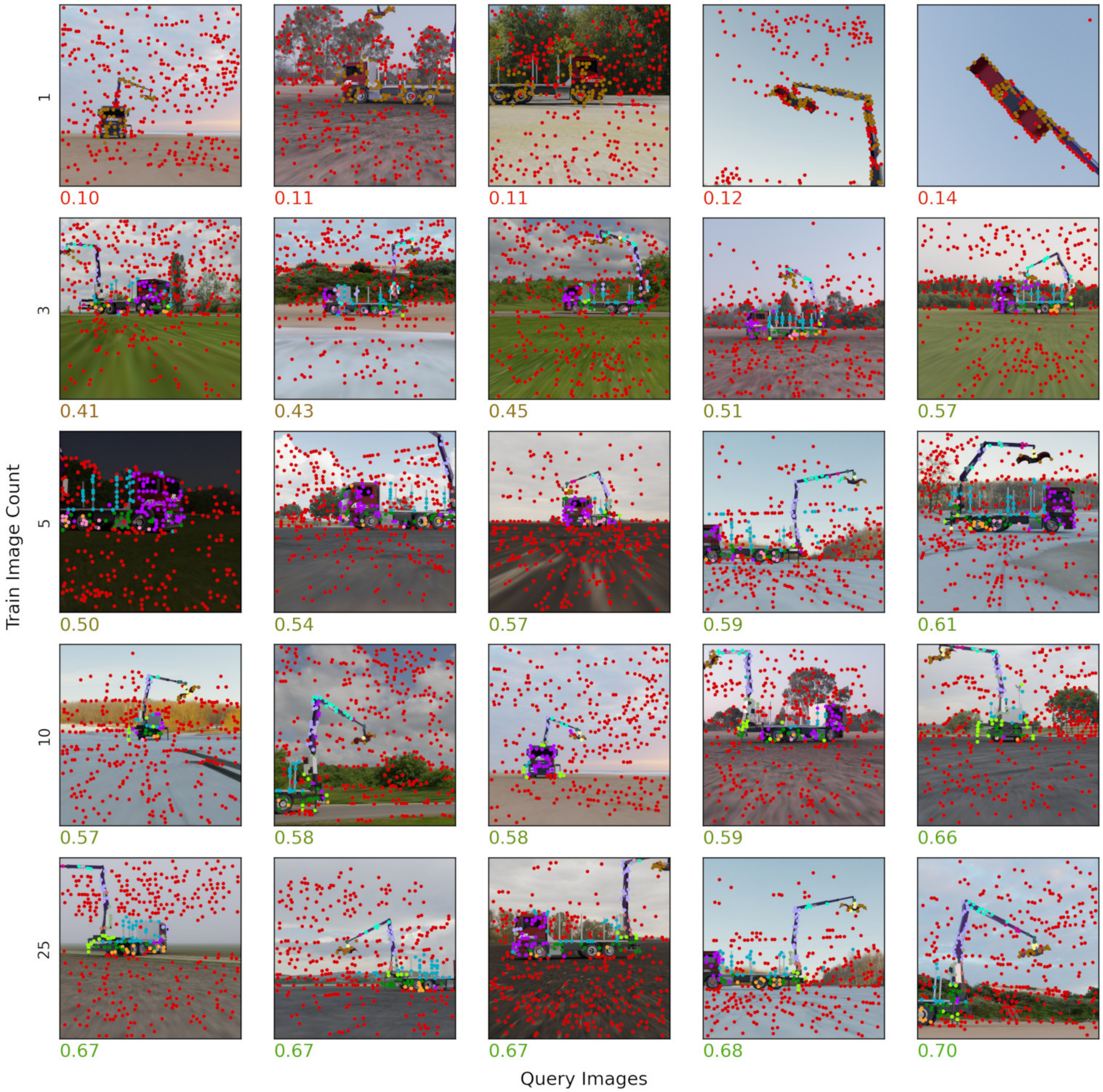

Figure 8. ***Medium*** granularity graph node classification results for different train sample sizes. Each row shows the best five classification results for the given train sample size measured with *F1* score.

Figure 9. ***Medium*** granularity semantic segmentation results for different train sample sizes. Each row shows the best five segmentation results for the given train sample size measured with *dice* score.

F. Samples: *High*

Figure 10. ***High*** granularity graph node classification results for different train sample sizes. Each row shows the best five classification results for the given train sample size measured with *F1* score.

Figure 11. ***High*** granularity semantic segmentation results for different train sample sizes. Each row shows the best five segmentation results for the given train sample size measured with *dice* score.

G. Mask R-CNN comparison

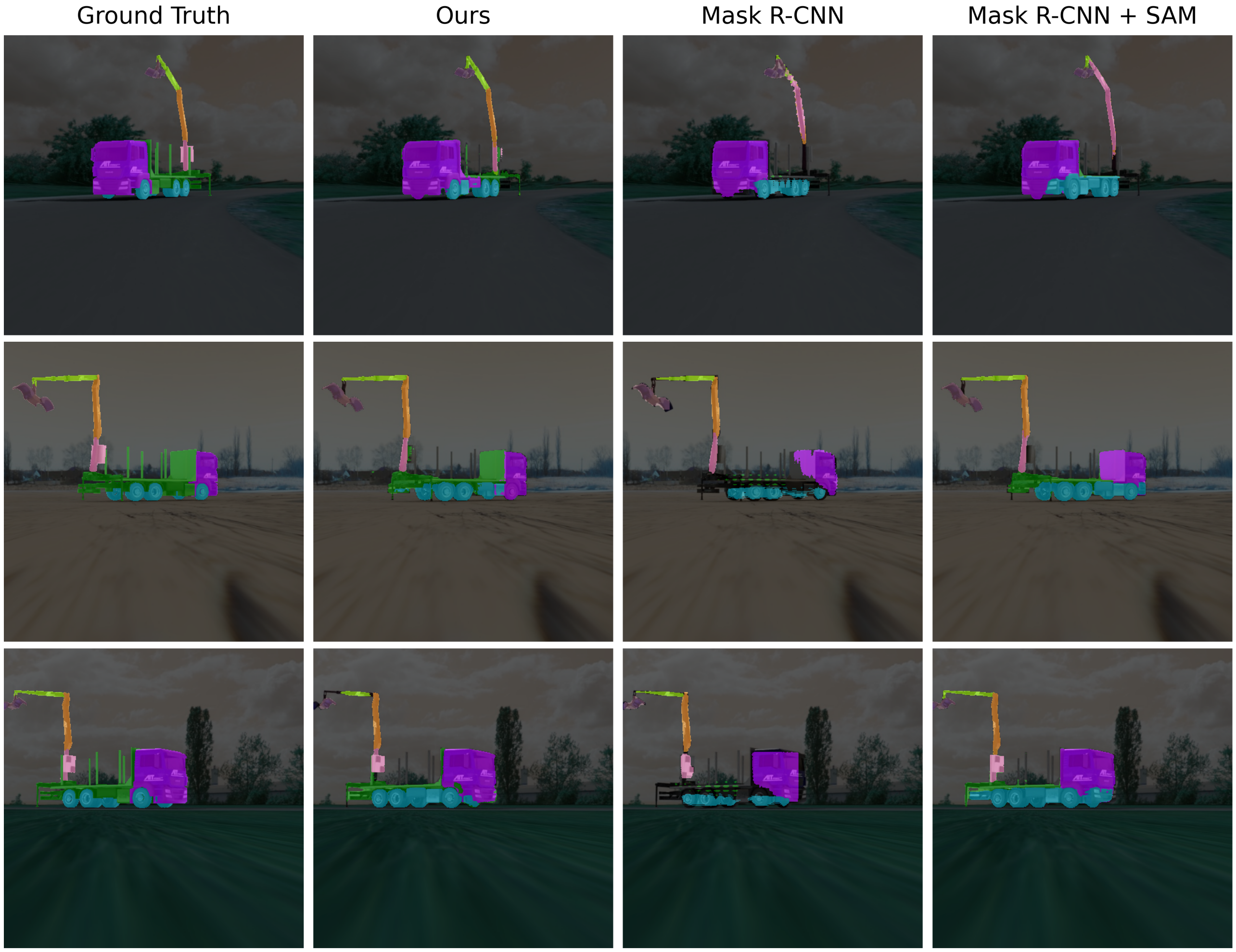

Figure 12. Our method, trained on just 10 samples, demonstrates competitive performance when compared to fine-tuning pre-trained(COCO) models like Mask R-CNN and Mask R-CNN + SAM (which uses estimated bounding boxes as SAM prompts). While Mask R-CNN was trained with 15 training and 10 validation samples, our method achieved superior results with fewer training data. This highlights the robustness of our approach, especially in low-data scenarios, where conventional methods like Mask R-CNN typically require larger datasets to perform optimally.