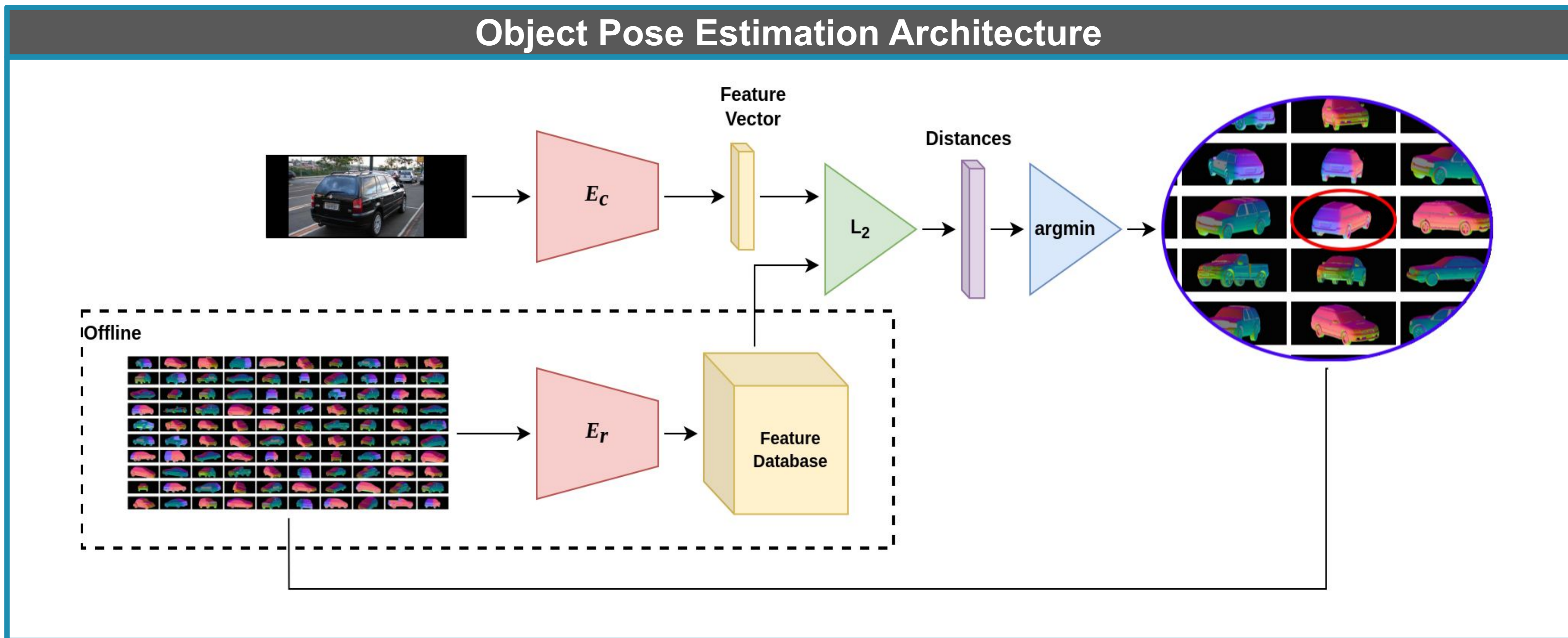




# Category-Level Pose Retrieval with Contrastive Features Learnt with Occlusion Augmentation



### Contrastive Pose Learning

- ❖ Loss is proportional to the **geodesic distance  $\Delta\theta$**  between a pair of samples
- ❖ The encoders  $E_c$  and  $E_r$  are trained jointly to **compare real objects to renderings**

### Data Augmentations

- ❖ From top-to-bottom: **Horizontal flipping, gaussian blurring, color jittering, bounding box augmentation, synthetic occlusion augmentation**
- ❖ Synthetic occlusions during training help the model learn to ignore foreground occlusions and background clutter

### Results

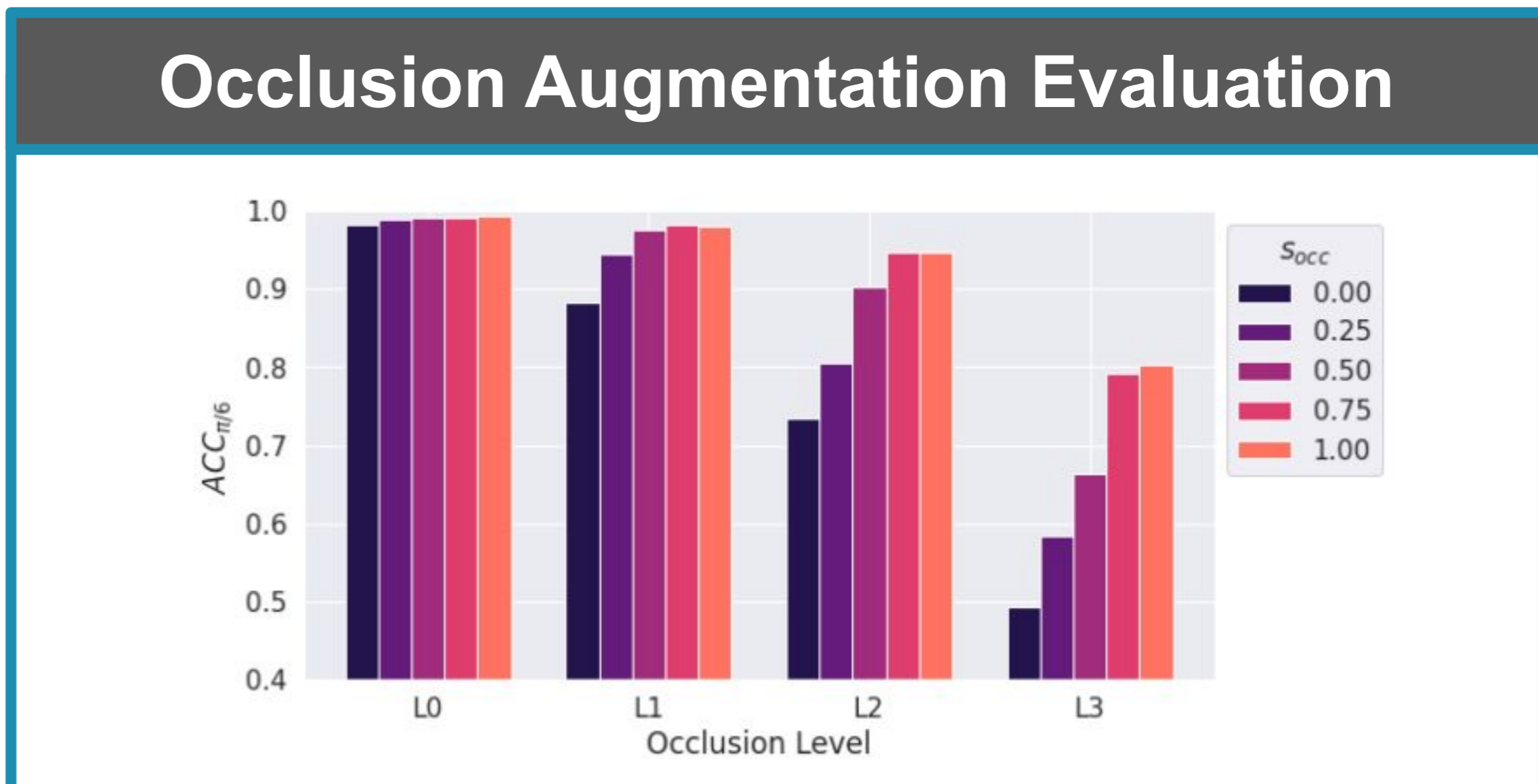
- ❖ PASCAL3D (L0) and OccludedPASCAL3D (L1-L3)

	Categ. aware	$ACC_{\frac{\pi}{6}} \uparrow$				$ACC_{\frac{\pi}{18}} \uparrow$				$MedErr \downarrow$			
		L0	L1	L2	L3	L0	L1	L2	L3	L0	L1	L2	L3
Res50-A <sup>†</sup>		88.1	70.4	52.8	37.8	44.6	25.3	14.5	6.7	11.7	17.9	30.4	46.4
Res50-S <sup>†</sup>	✓	87.6	73.2	58.4	43.1	43.9	28.1	18.6	9.9	11.8	17.3	26.1	44.0
StarMap <sup>†</sup>		89.4	71.1	47.2	22.9	59.5	34.4	13.9	3.7	9.0	17.6	34.1	63.0
NeMo <sup>†</sup>	✓	84.1	73.1	59.9	41.3	60.4	45.1	30.2	14.5	9.3	15.6	24.1	41.8
NeMo-M <sup>†</sup>	✓	86.7	77.2	65.2	47.1	63.2	49.9	34.5	17.8	8.2	13.0	20.2	<b>36.1</b>
NeMo-S <sup>†</sup>	✓	86.1	76.0	63.9	46.8	61.0	46.3	32.0	17.1	8.8	13.6	20.2	36.5
PoseCon		90.8	76.2	59.3	39.7	67.2	46.4	28.1	12.7	7.1	12.6	23.1	45.5
Ours	✓	<b>92.3</b>	<b>85.7</b>	<b>72.7</b>	<b>49.8</b>	<b>72.2</b>	<b>56.7</b>	<b>38.9</b>	<b>17.9</b>	<b>6.6</b>	<b>9.7</b>	<b>16.0</b>	37.9

- ❖ KITTI3D (Fully Visible, Partly Occluded, Largely Occluded)

	$ACC_{\frac{\pi}{6}} \uparrow$				$ACC_{\frac{\pi}{18}} \uparrow$				$MedErr \downarrow$			
	FV	PO	LO	All	FV	PO	LO	All	FV	PO	LO	All
NeMo-S	88.1	72.4	34.9	67.9	70.3	40.4	7.5	43.7	7.3	11.6	46.1	20.0
PoseContrast	97.8	88.5	48.6	80.6	81.6	62.4	18.6	57.5	6.6	8.6	33.0	15.0
Ours	<b>98.1</b>	<b>90.0</b>	<b>56.1</b>	<b>83.4</b>	<b>92.8</b>	<b>70.6</b>	<b>21.0</b>	<b>65.3</b>	<b>3.2</b>	<b>5.4</b>	<b>24.8</b>	<b>10.2</b>
Ours-2	97.9	<b>90.6</b>	<b>66.5</b>	<b>86.5</b>	<b>94.2</b>	<b>74.4</b>	<b>34.4</b>	<b>70.9</b>	<b>2.9</b>	<b>5.3</b>	<b>15.5</b>	<b>7.3</b>

(Ours-2 is trained with stronger bounding box augmentation)



### Acknowledgements

This work has been supported by the Sim2Real project, in the context of the Ford-KU Leuven alliance program.