

机器人智能视觉感知

鲁继文

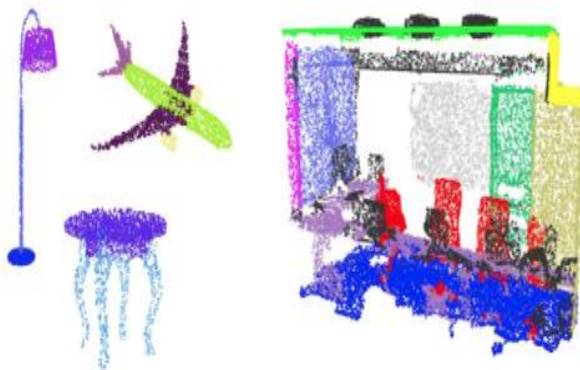
清华大学自动化系

2020年7月28日

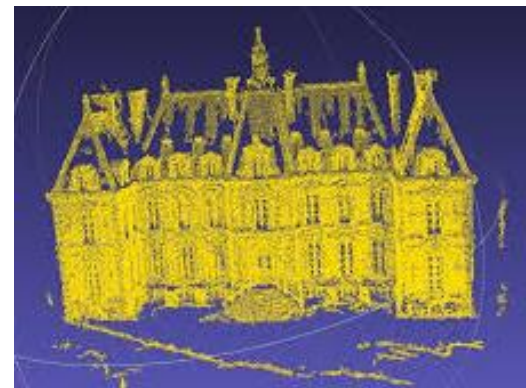
• 机器人视觉感知



单目三维检测



三维点云识别



三维点云重建



三维布局估计



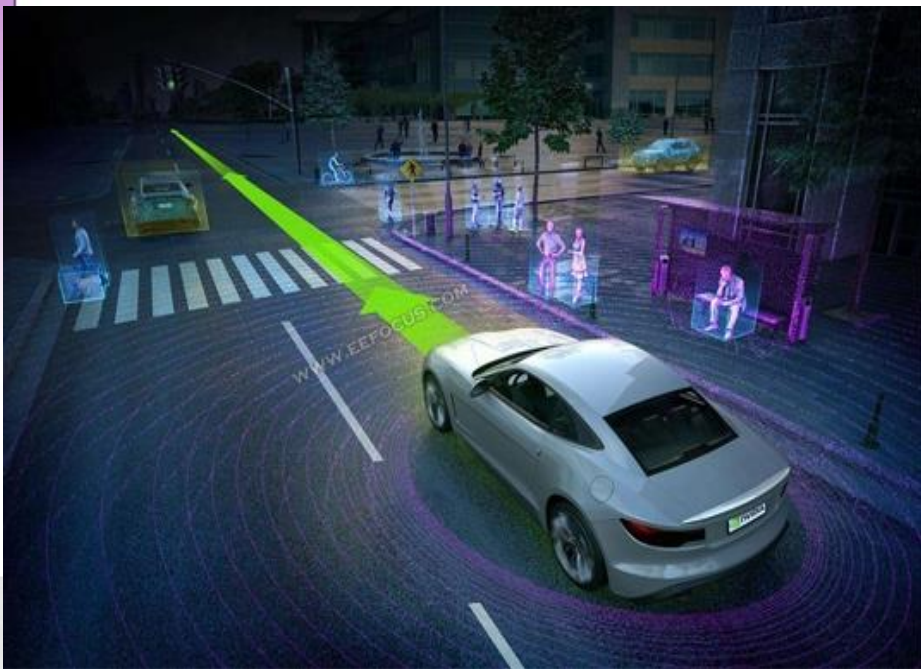
机器人边缘计算



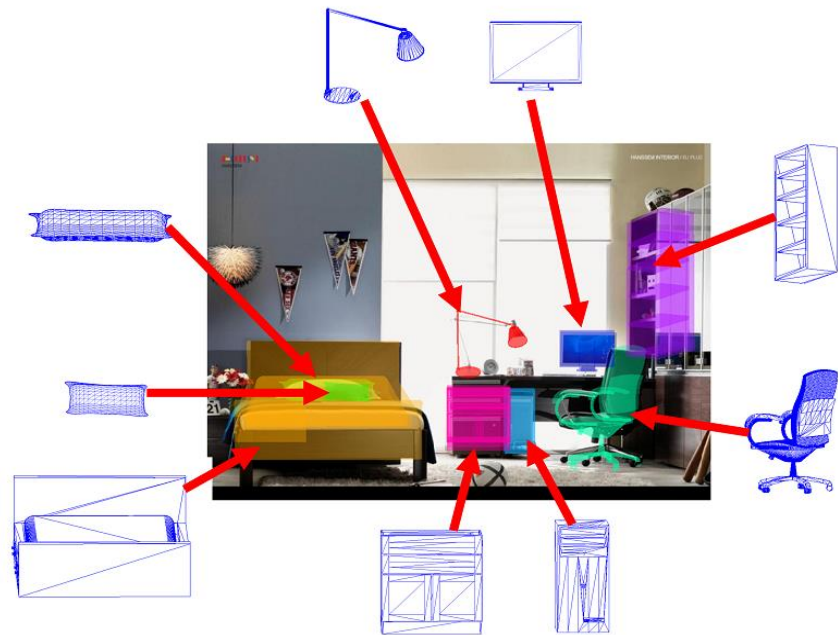
机器人目标抓取

• 机器人视觉感知

➤ 三维环境感知



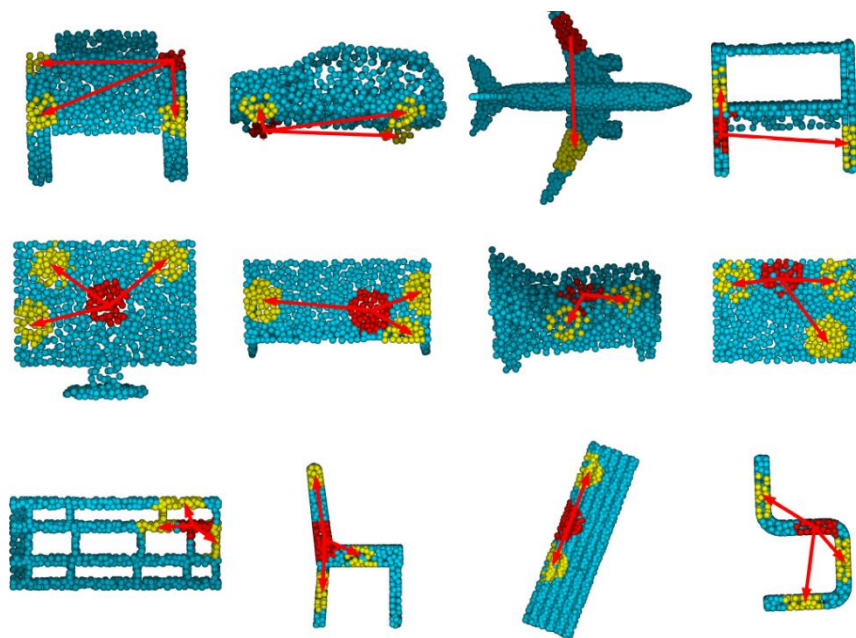
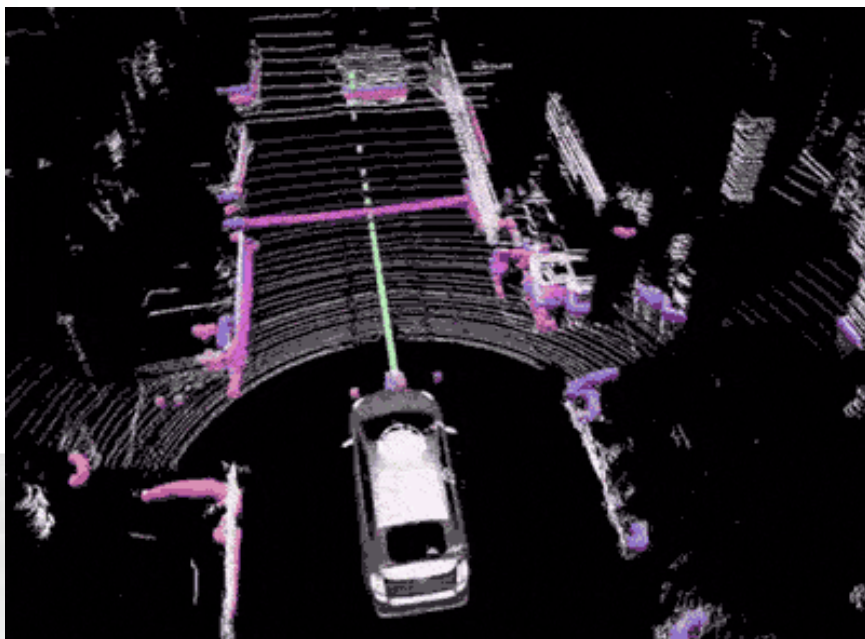
室外环境：行人、车辆



室内环境：家具、布局

• 三维环境感知

- 三维点云被广泛应用于机器人视觉感知任务，比传统图像具有更便捷的信息获取方式

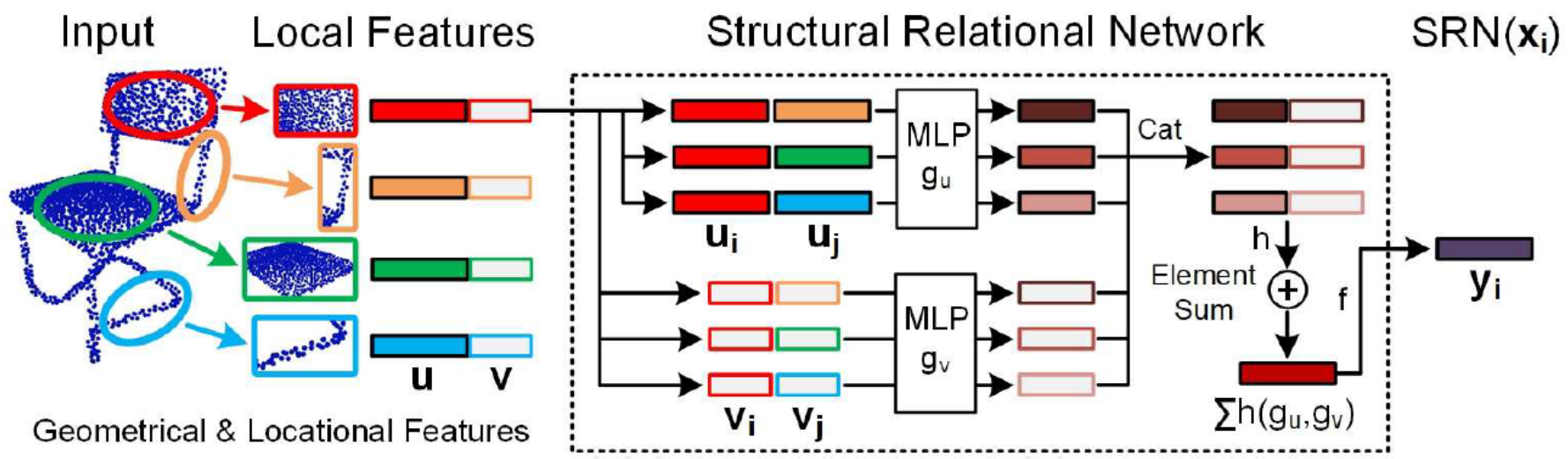


• 我们的工作

➤ 三维环境感知

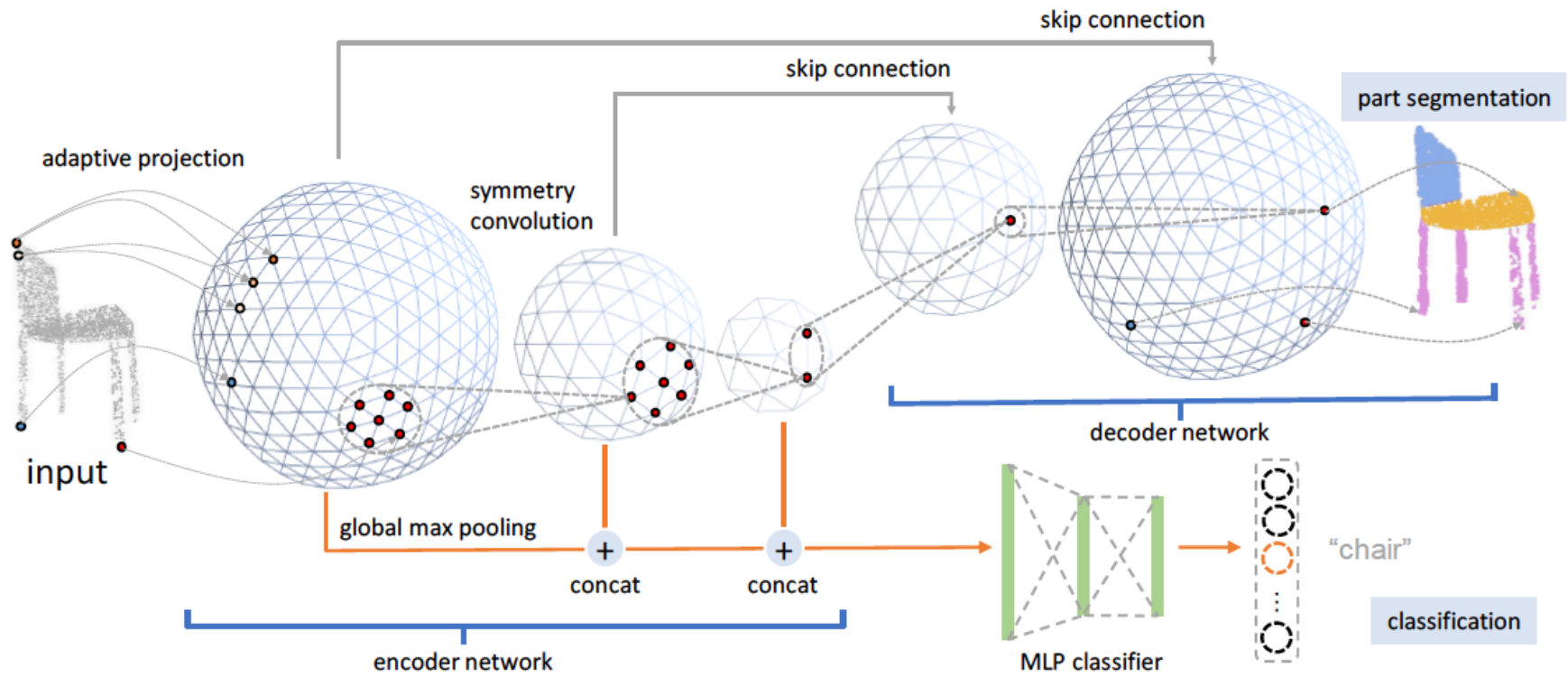
- Structural Relational Reasoning of Point Clouds, *CVPR*, 2019
- Spherical Fractal Convolution Neural Networks for Point Cloud Recognition, *CVPR*, 2019
- Conditional Single-view Shape Generation for Multi-view Stereo Reconstruction, *CVPR*, 2019
- Deep Fitting Degree Scoring Network for Monocular 3D Object Detection, *CVPR*, 2019
- Global-Local Bidirectional Reasoning for Unsupervised Representation Learning of 3D Point Clouds, *CVPR*, 2020
- Reinforced Axial Refinement Network for Monocular 3D Object Detection, *ECCV*, 2020
- Rotation-robust Intersection over Union for 3D Object Detection, *ECCV*, 2020
- Structural Deep Metric Learning for Room Layout Estimation, *ECCV*, 2020
- Spatial Geometric Reasoning for Room Layout Estimation via Deep Reinforcement Learning, *ECCV*, 2020

• 基于结构关系推理的点云分割与识别



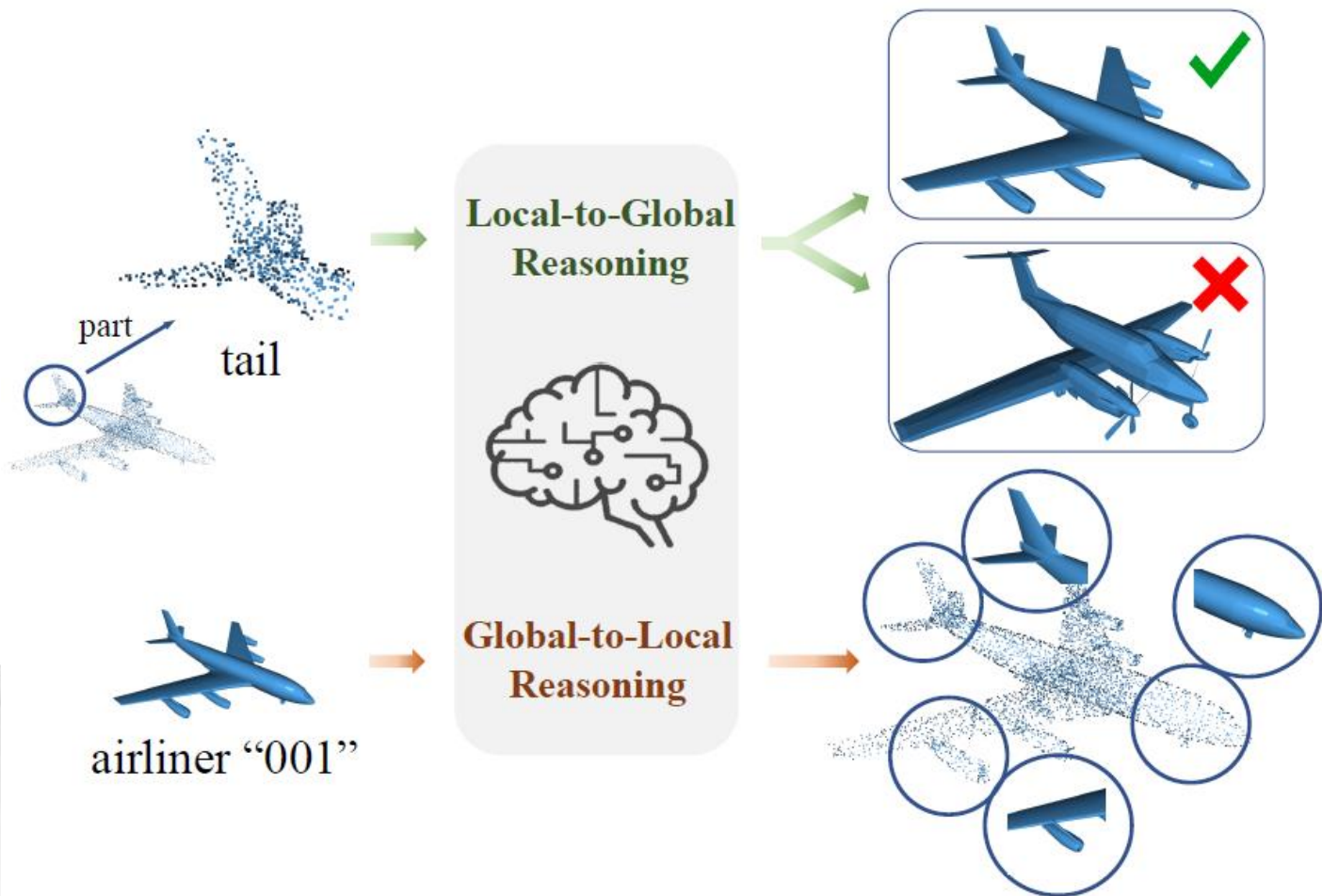
[1] Yueqi Duan, Yu Zheng, **Jiwen Lu**, Jie Zhou, and Qi Tian, Structural Relational Reasoning of Point Clouds, *CVPR*, pp. 949-958, 2019.

• 基于球面分形网络的点云目标识别

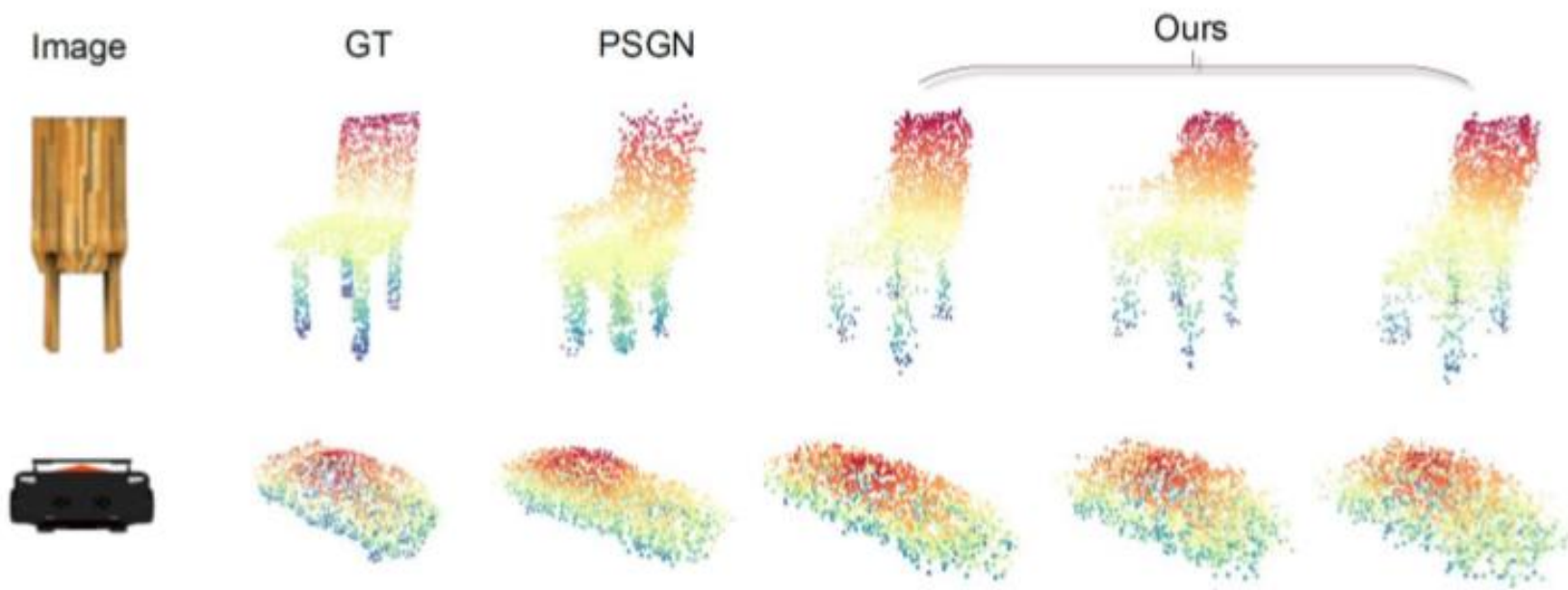


[2] Yongming Rao, **Jiwen Lu**, and Jie Zhou, Spherical Fractal Convolutional Neural Networks for Point Cloud Recognition, *CVPR*, pp. 452-460, 2019.

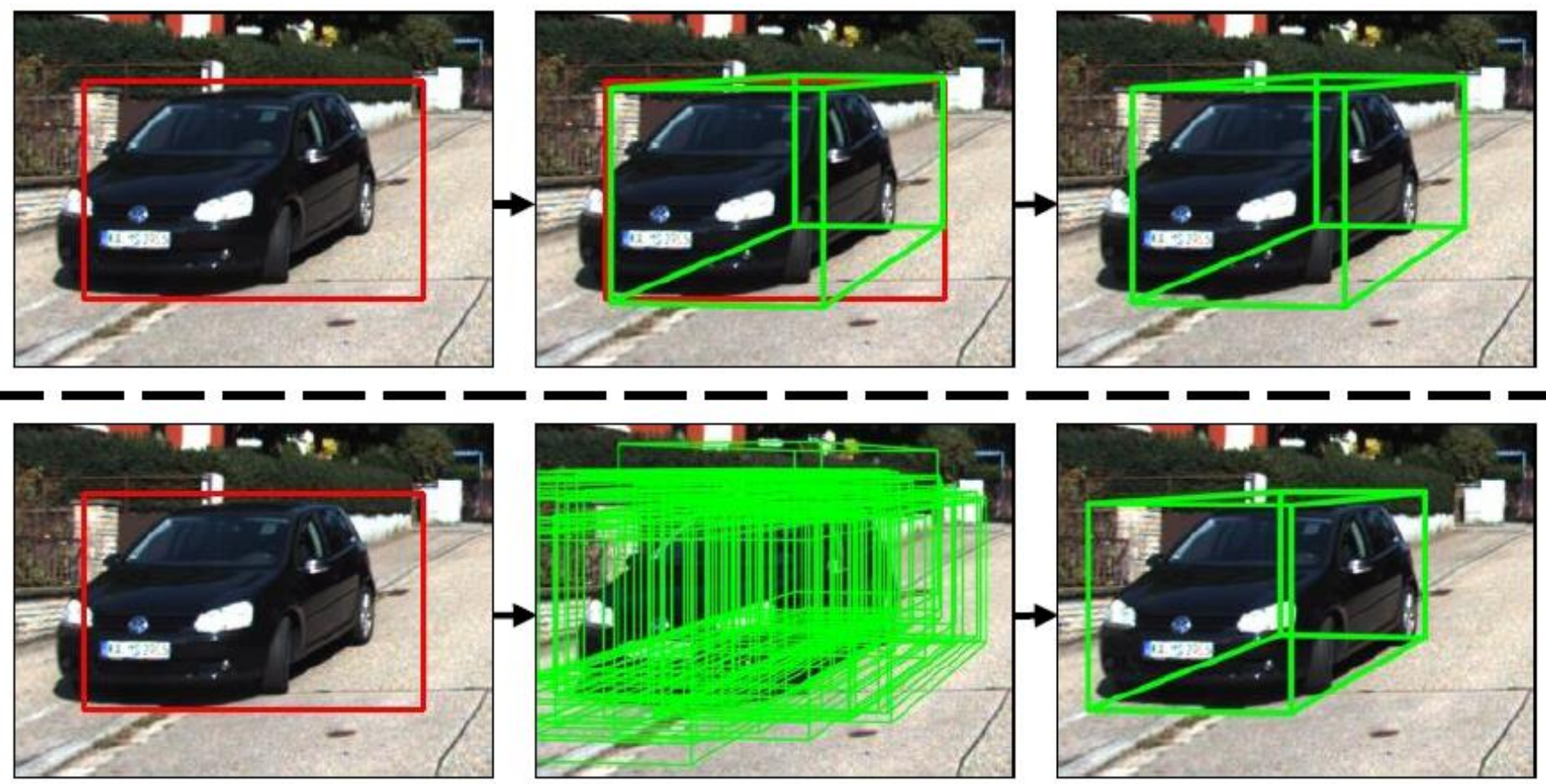
• 基于双向推理表示的点云目标识别



• 基于一致性约束的多视图立体重建



• 基于深度拟合网络的单目三维检测

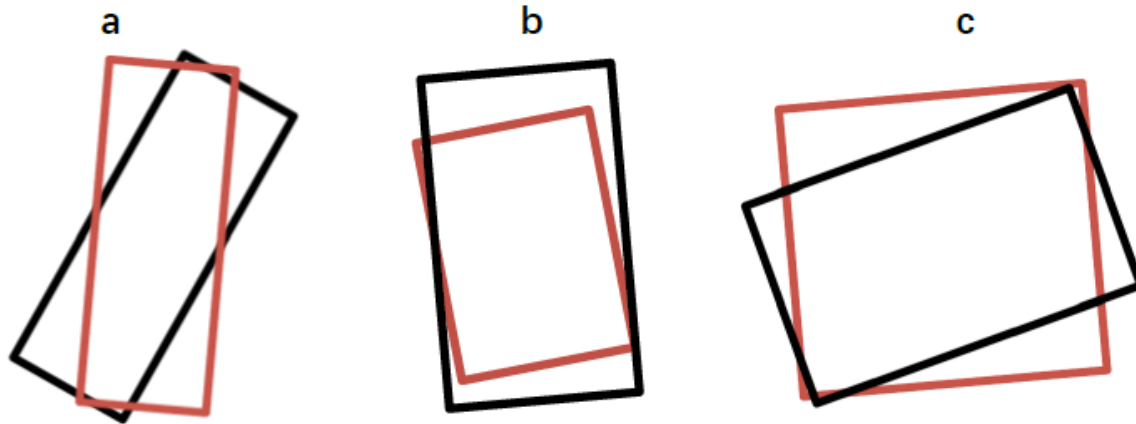


• 基于深度强化学习的单目三维检测



[6] Lijie Liu, Chufan Wu, **Jiwen Lu**, Lingxi Xie, Jie Zhou, and Qi Tian, Reinforced Axial Refinement Network for Monocular 3D Object Detection, *ECCV*, 2020, accepted.

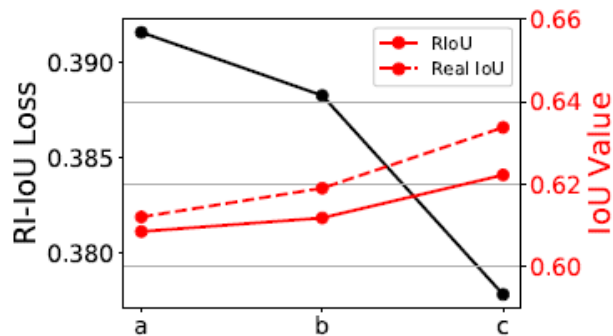
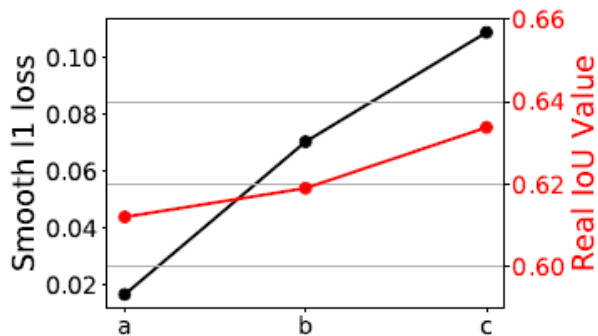
• 基于旋转鲁棒交并比的三维目标检测



Smooth L1 loss = 0.0166
Real IoU = 0.6120
RI-IoU = 0.6085
RI-IoU Loss = 0.3915

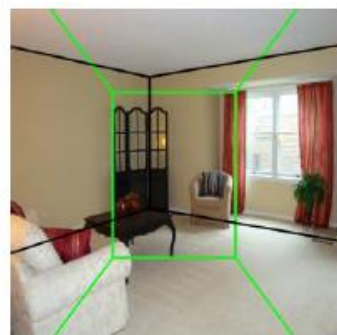
Smooth L1 loss = 0.0703
Real IoU = 0.6190
RI-IoU = 0.6118
RI-IoU Loss = 0.3882

Smooth L1 loss = 0.1087
Real IoU = 0.6337
RI-IoU = 0.6222
RI-IoU Loss = 0.3778



[7] Yu Zheng, Danyang Zhang, Sinan Xie, **Jiwen Lu**, and Jie Zhou, Rotation-robust Intersection over Union for 3D Object Detection, *ECCV*, 2020, accepted.

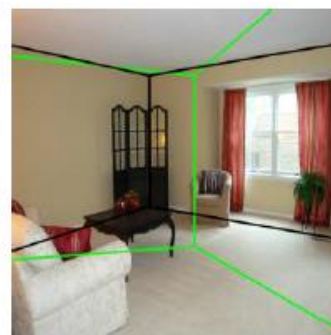
• 基于深度强化学习的三维布局估计



o_1

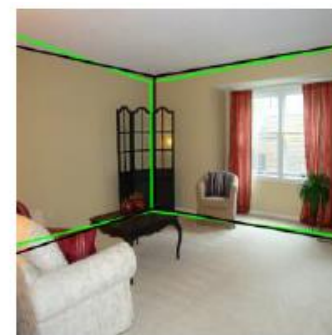


o_2



o_3

...



o_{T+1}



s_1



a_1



s_2



a_2



s_3



...



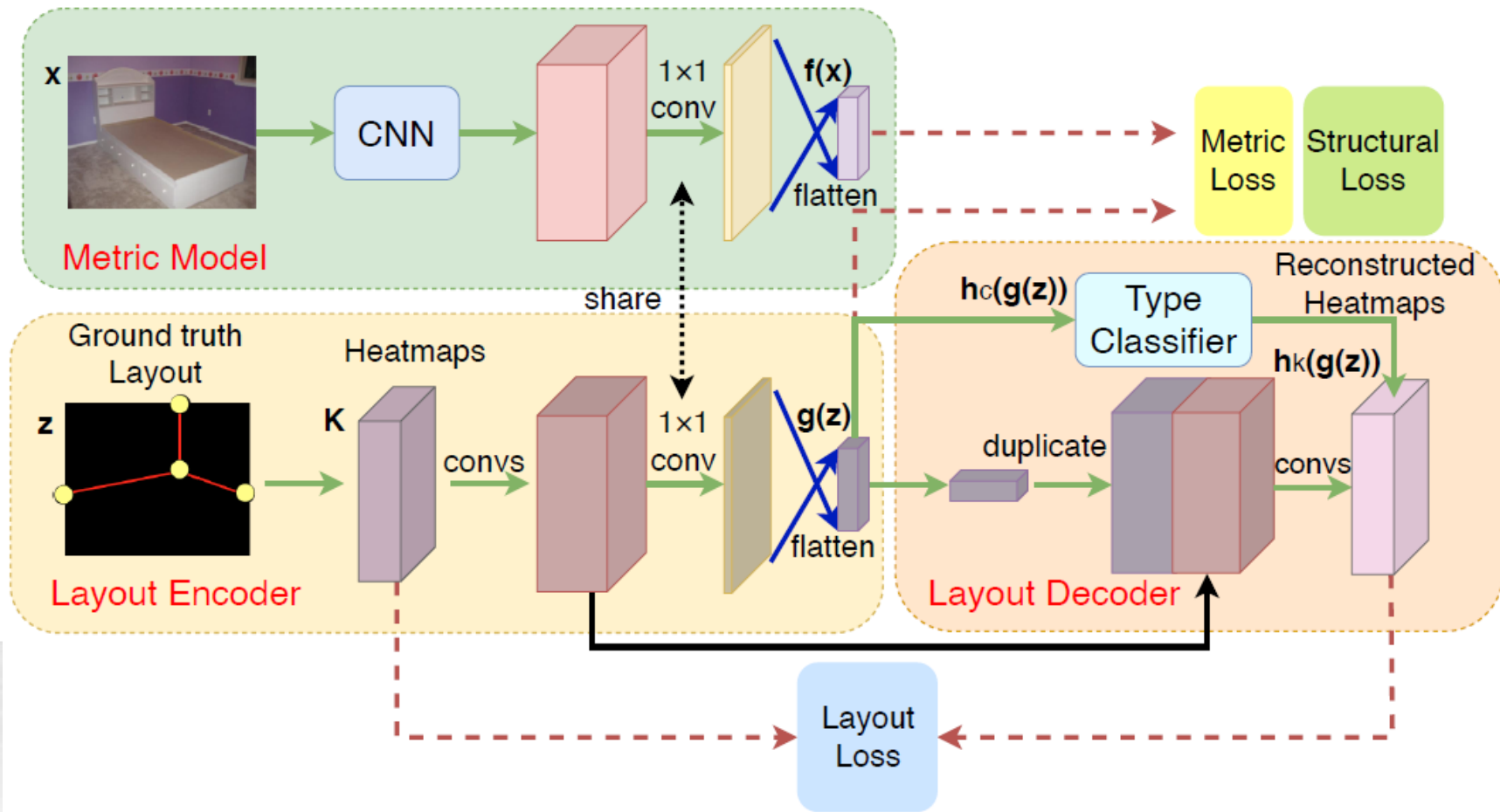
a_T



s_{T+1}

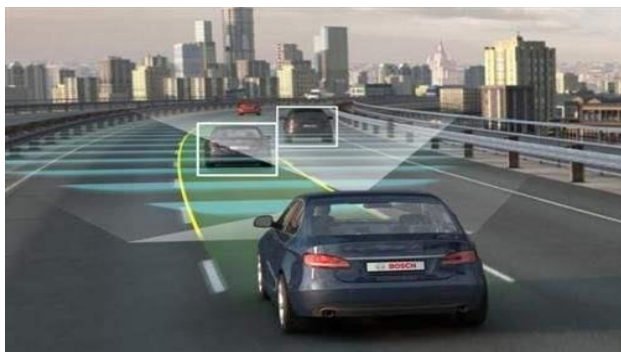
[8] Liangliang Ren, Yangyang Song, **Jiwen Lu**, and Jie Zhou, Spatial Geometric Reasoning for Room Layout Estimation via Deep Reinforcement Learning, *ECCV*, 2020, accepted.

• 基于结构度量学习的三维布局估计

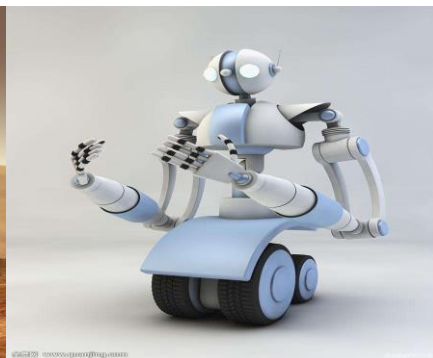


• 机器人边缘计算

➤ 深度模型参数多计算量大，无法直接部署



无人驾驶



机器人



可穿戴设备

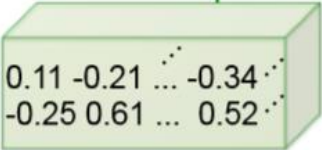
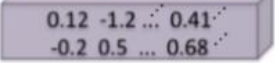
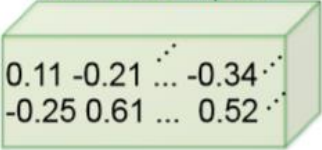
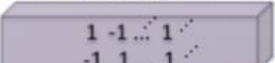
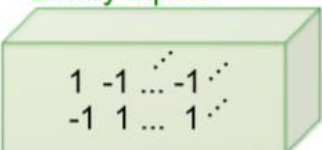
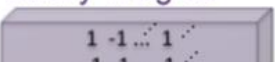


智能手机



• 机器人边缘计算

➤ 减少计算量与存储量，保持云计算的性能

	Network Variations	Operations used in Convolution	Memory Saving (Inference)	Computation Saving (Inference)
Standard Convolution	<p>Real-Value Inputs</p>  <p>Real-Value Weights</p> 	$+, -, \times$	1x	1x
Binary Weight	<p>Real-Value Inputs</p>  <p>Binary Weights</p> 	$+, -$	$\sim 32x$	$\sim 2x$
BinaryWeight Binary Input (XNOR-Net)	<p>Binary Inputs</p>  <p>Binary Weights</p> 	XNOR, bitcount	$\sim 32x$	$\sim 58x$

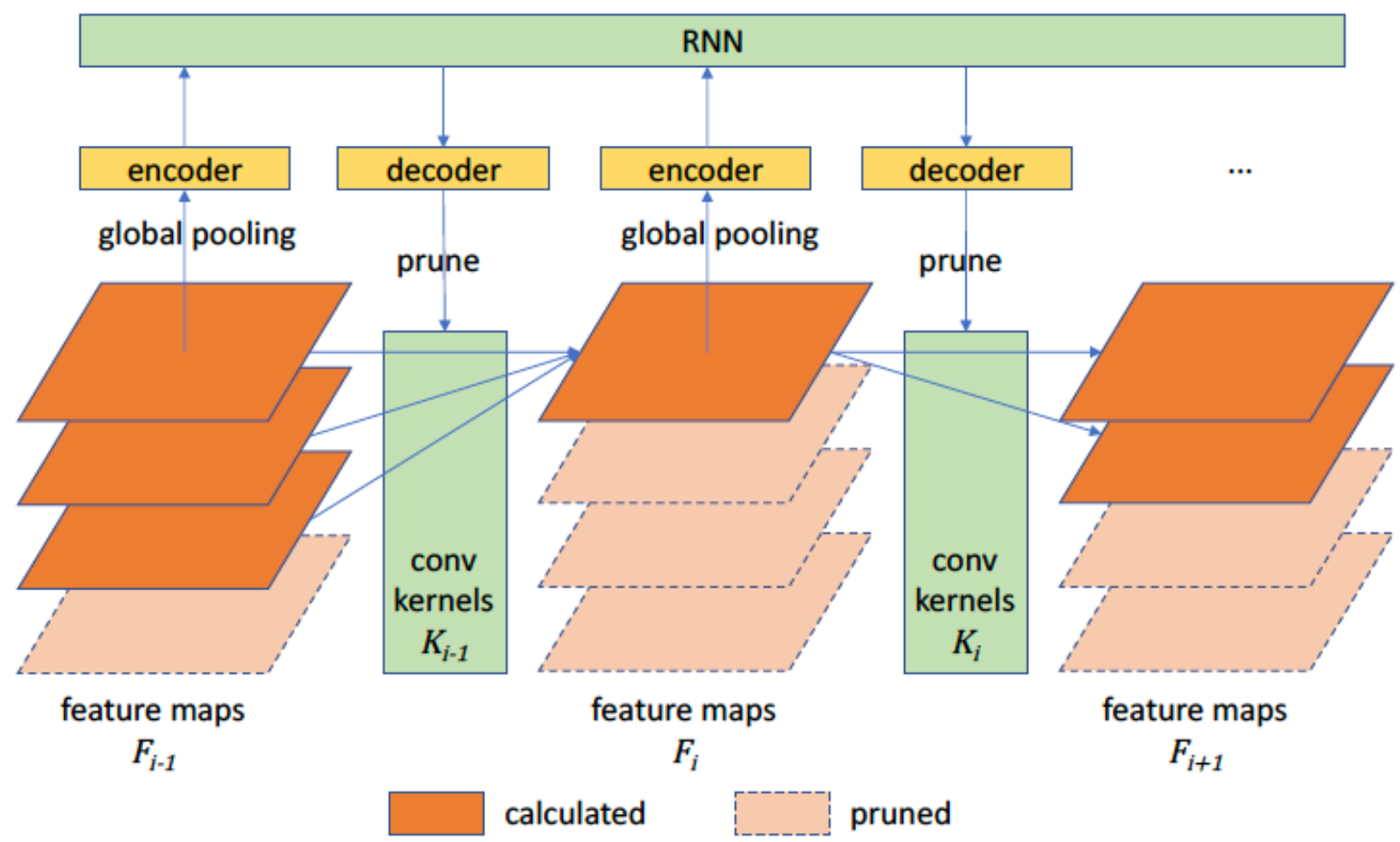
• 我们的工作

➤ 机器人边缘计算

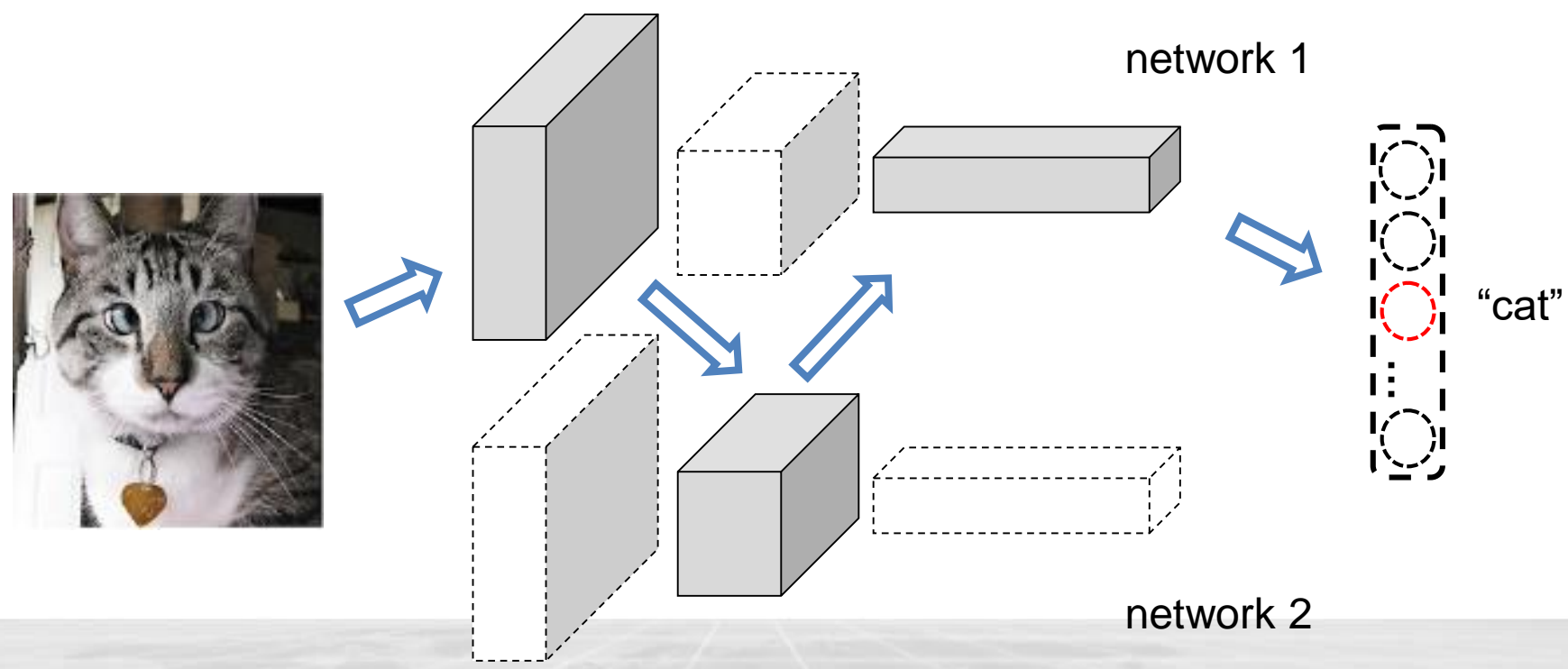
- Runtime Neural Pruning, *NeurIPS*, 2017
- Learning Channel-wise Interactions for Binary Convolutional Neural Networks, *CVPR*, 2019
- Enhanced Bayesian Compression via Deep Reinforcement Learning, *CVPR*, 2019
- BiDet: An Efficient Binarized Object Detector, *CVPR*, 2020
- MetaDistiller: Network Self-boosting via Meta-learned Top-down Distillation, *ECCV*, 2020
- Runtime Network Routing for Efficient Image Classification, *PAMI*, 2019
- Learning Channel-wise Interactions for Binary Convolutional Neural Networks, *PAMI*, 2020



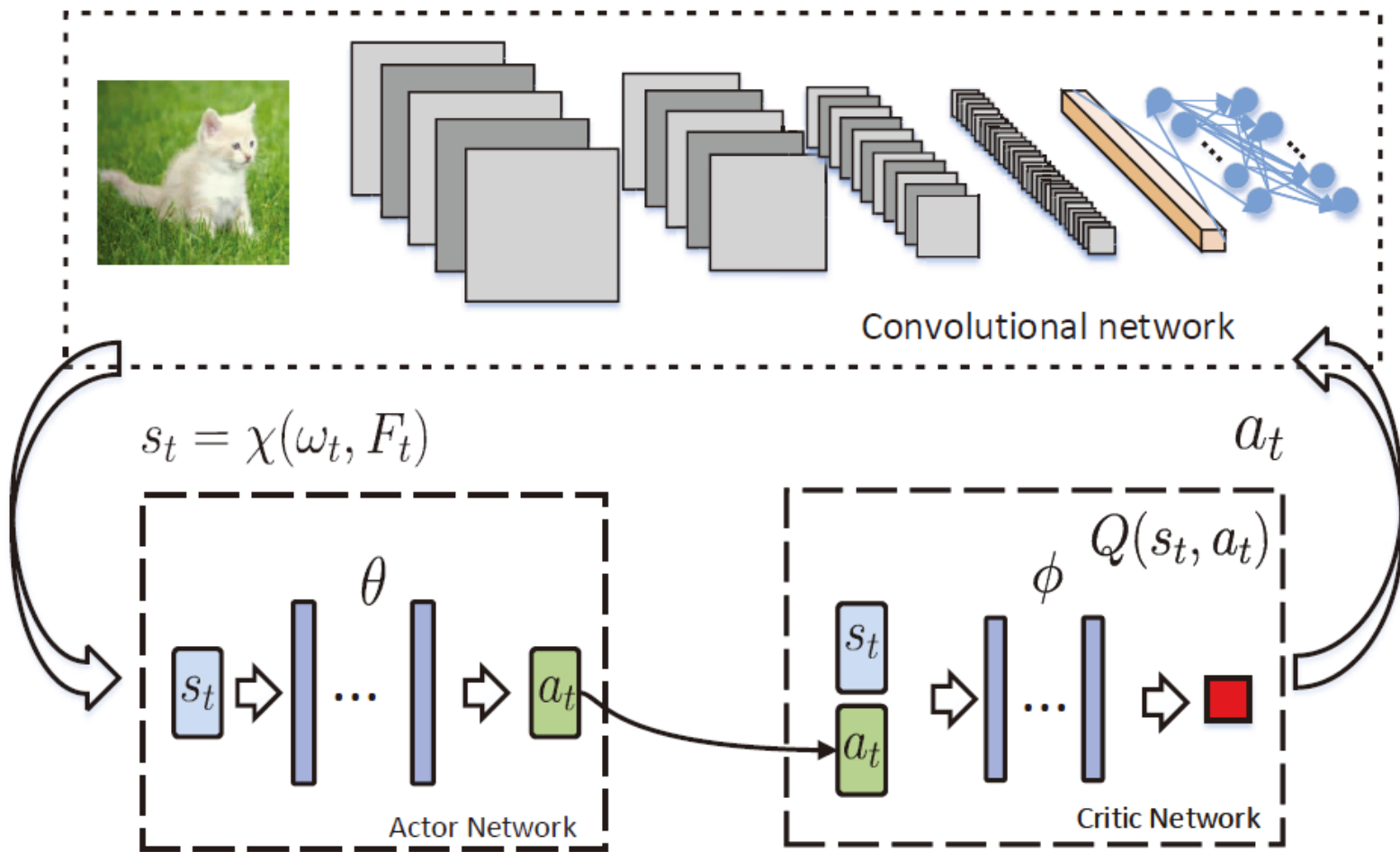
• 基于动态神经网络的深度模型压缩



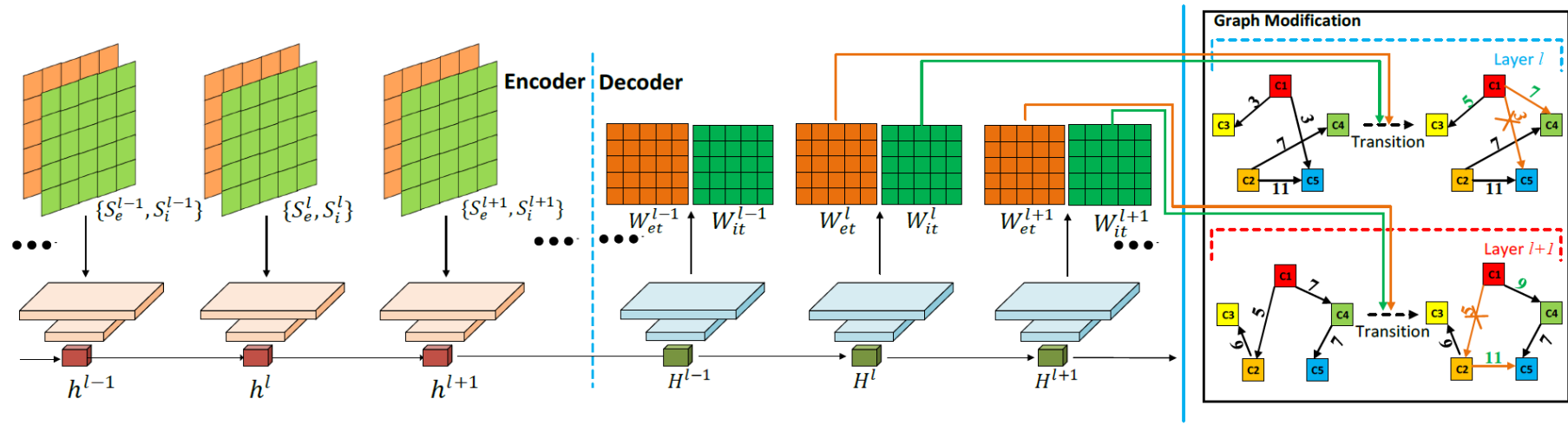
• 基于动态神经网络的深度模型压缩



• 基于深度强化学习的贝叶斯模型压缩

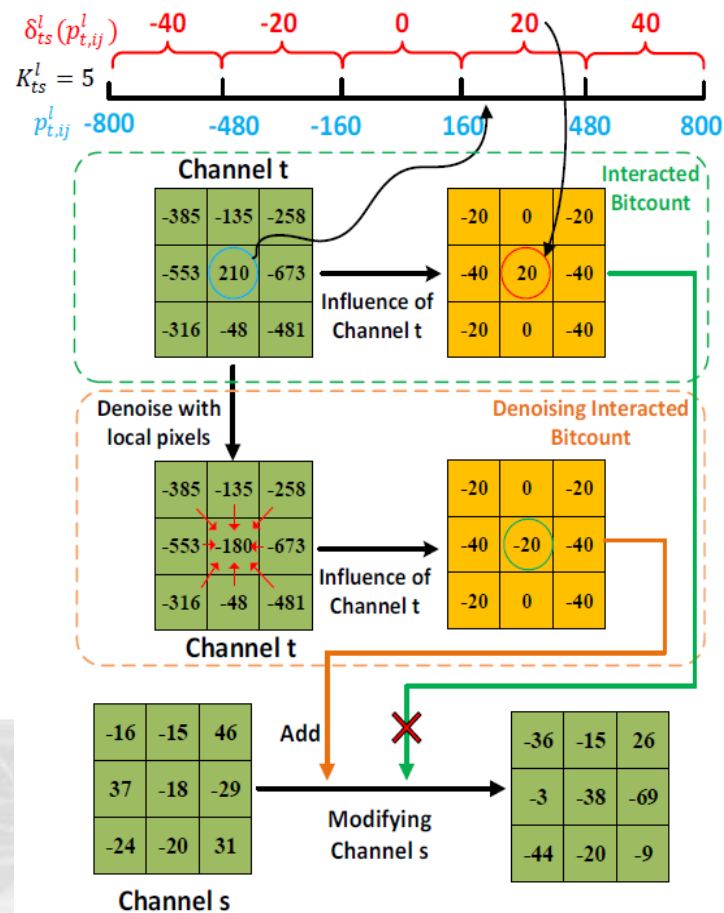
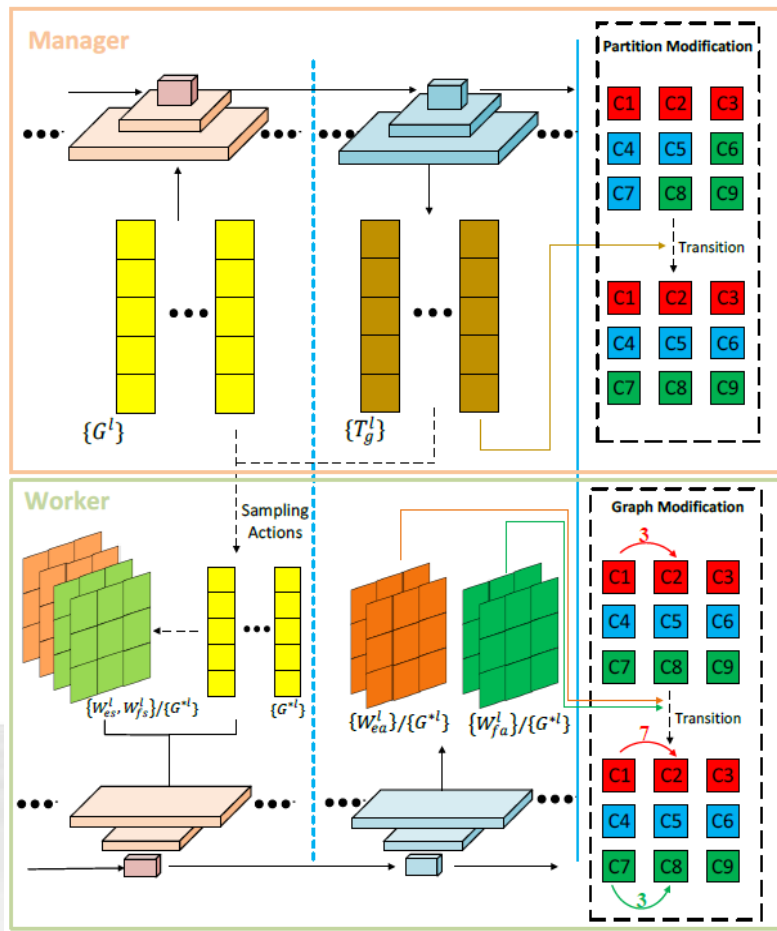


• 基于通道交互的二值卷积神经网络

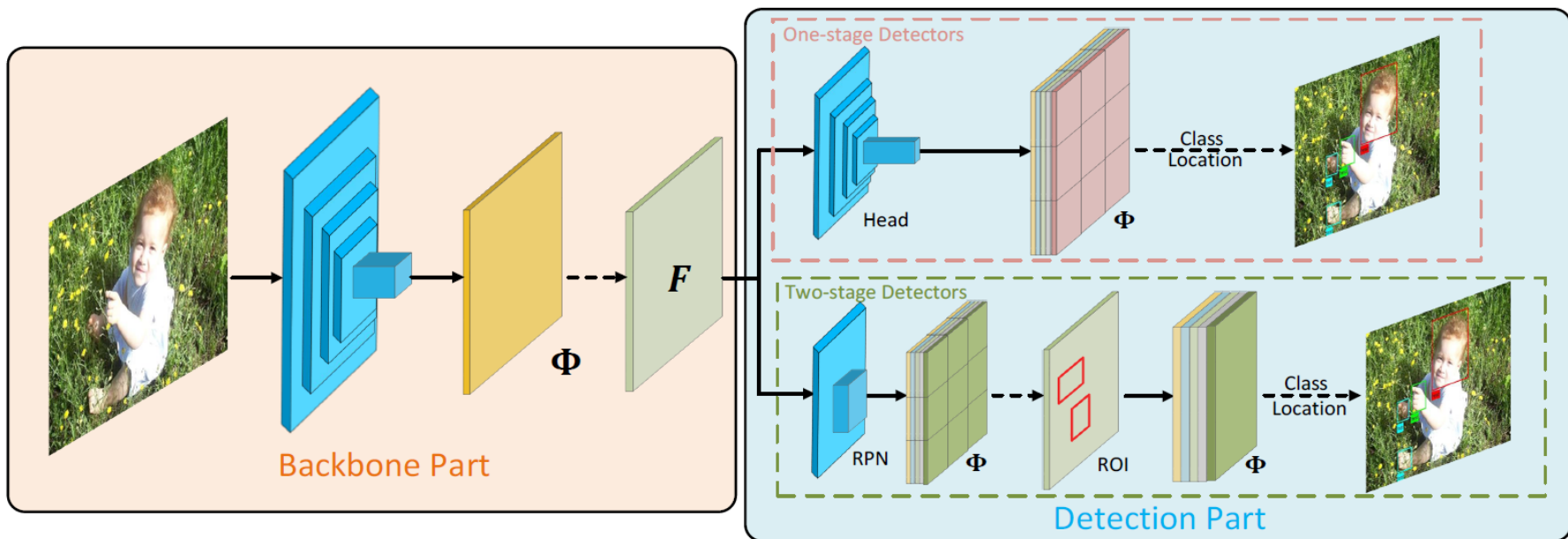


[13] Ziwei Wang, **Jiwen Lu**, Chenxin Tao, Jie Zhou, and Qi Tian, Learning Channel-wise Interactions for Binary Convolutional Neural Networks, *CVPR*, pp. 568-577, 2019.

• 基于通道交互的二值卷积神经网络



• 基于二值神经网络的视觉目标检测



[15] Ziwei Wang, Ziyi Wu, **Jiwen Lu**, and Jie Zhou, BiDet: An Efficient Binarized Object Detector, *CVPR*, pp. 2049-2058, 2020.

谢谢！