2016年9月

智能驾驶与视觉感知

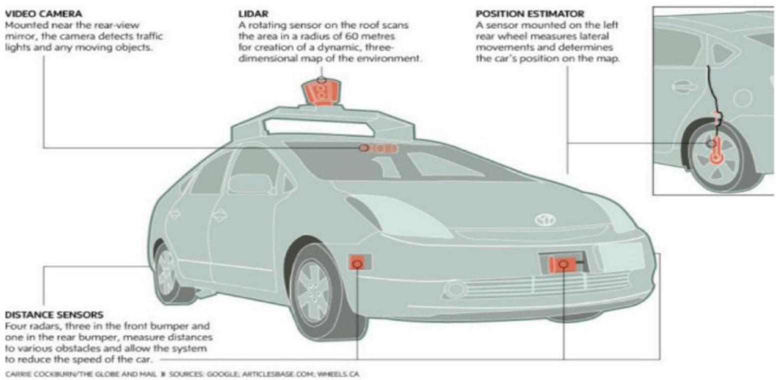
清华大学电子系　马惠敏

http://3dimage.ee.tsinghua.edu.cn　Email: mhmpub@tsinghua.edu.cn　Tel: 62781432

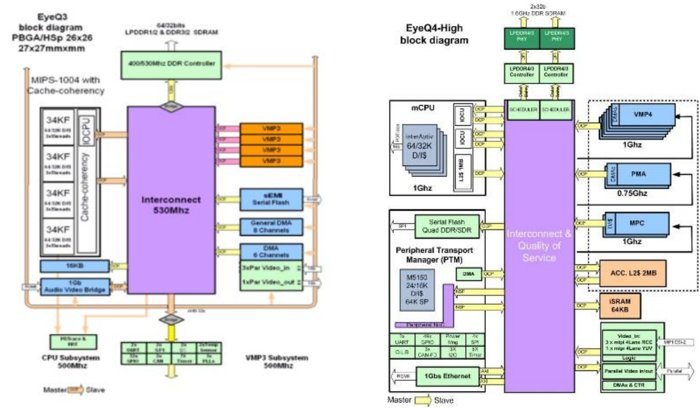
智能驾驶与视觉感知

Vision Conception for Semi-Autonomous Driving

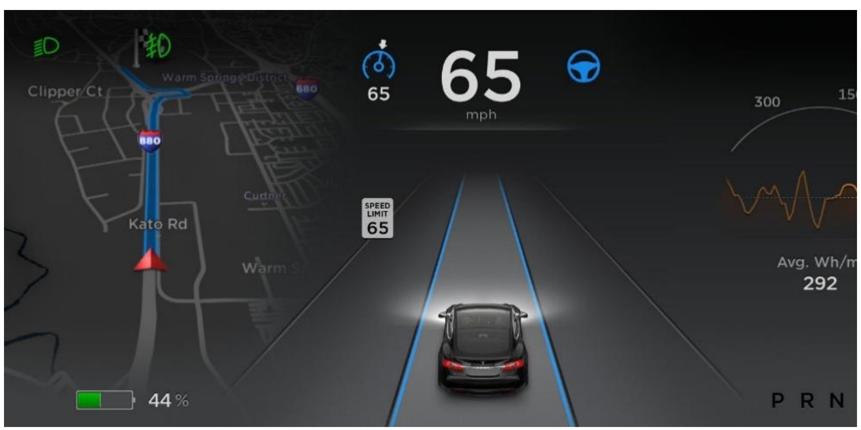
Google



Mobileye



Tesla



NVIDIA

清華大學
Tsinghua University

智能驾驶与视觉感知

Vision Conception for Semi-Autonomous Driving

Google

- LIDAR
- Radar/ Sonar
- Stereo Camera
- Maps

Mobileye

- EyeQ3 (Now)
- EyeQ4 (2018)

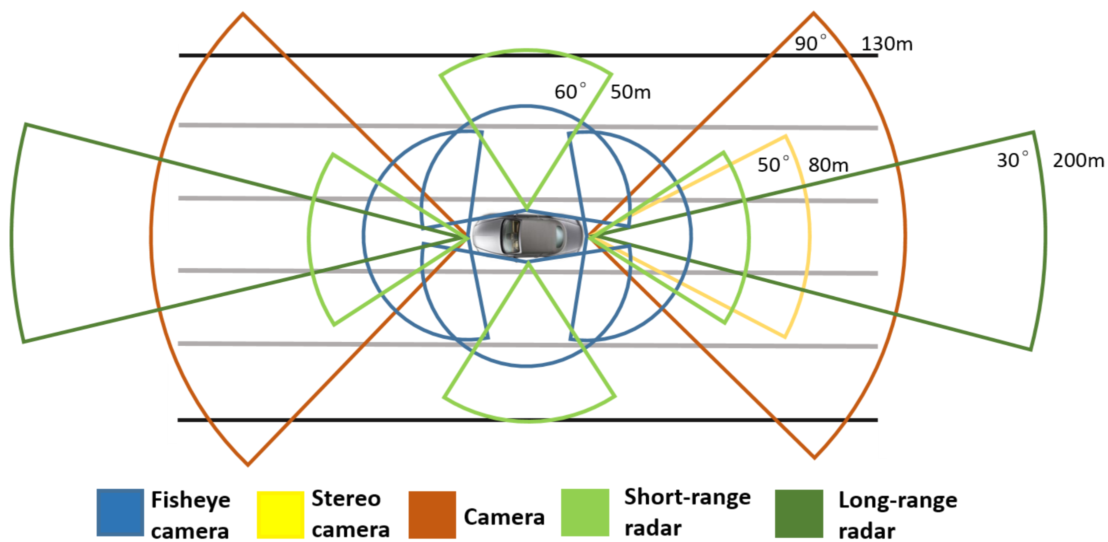
  8 cameras

NVIDIA

- Drive PX： 2 Jetson TX1,
  12 cameras
- Drive PX2： 12 CPU cores,
  2 Pascal GPUs
  DriveNet

Tesla

- Mobileye EyeQ3
- Forward Radar
- Forward Camera
- 360° Ultrasonic Sonar
- GPS



智能驾驶与视觉感知

Vision Conception for Semi-Autonomous Driving

□ 功能/Functions

- 前向碰撞预警
  Forward Collision Warning
- 自适应巡航控制
  Adaptive Cruise Control
- 自动变道
  Automatic Lane Changing
- 交通拥堵辅助
  Traffic Jam Assistance
- 自动泊车
  Autonomous parking

□ 技术/Techniques

- 车道检测
  Line Detection
- 交通标识/灯检测
  Traffic Sign/Light Detection
- 驾驶道路标识
  Drivable Space Labeling
- 通用物体检测
  General Object Detection
- 整体路径规划
  Holistic Path Planning

智能驾驶与视觉感知

Vision Conception for Semi-Autonomous Driving

核心问题：小目标、强遮挡、高动态

Small objects, Heavy Occlusion, Rapid moving

智能

ImageNet:

- 1000 classes, ~2M images


PASCAL VOC:

- 20 classes, ~20K images


Professional photographs, simple scenes, large objects recall> **95%** (1K Prop.)

KITTI:

- 3 classes, ~20K images



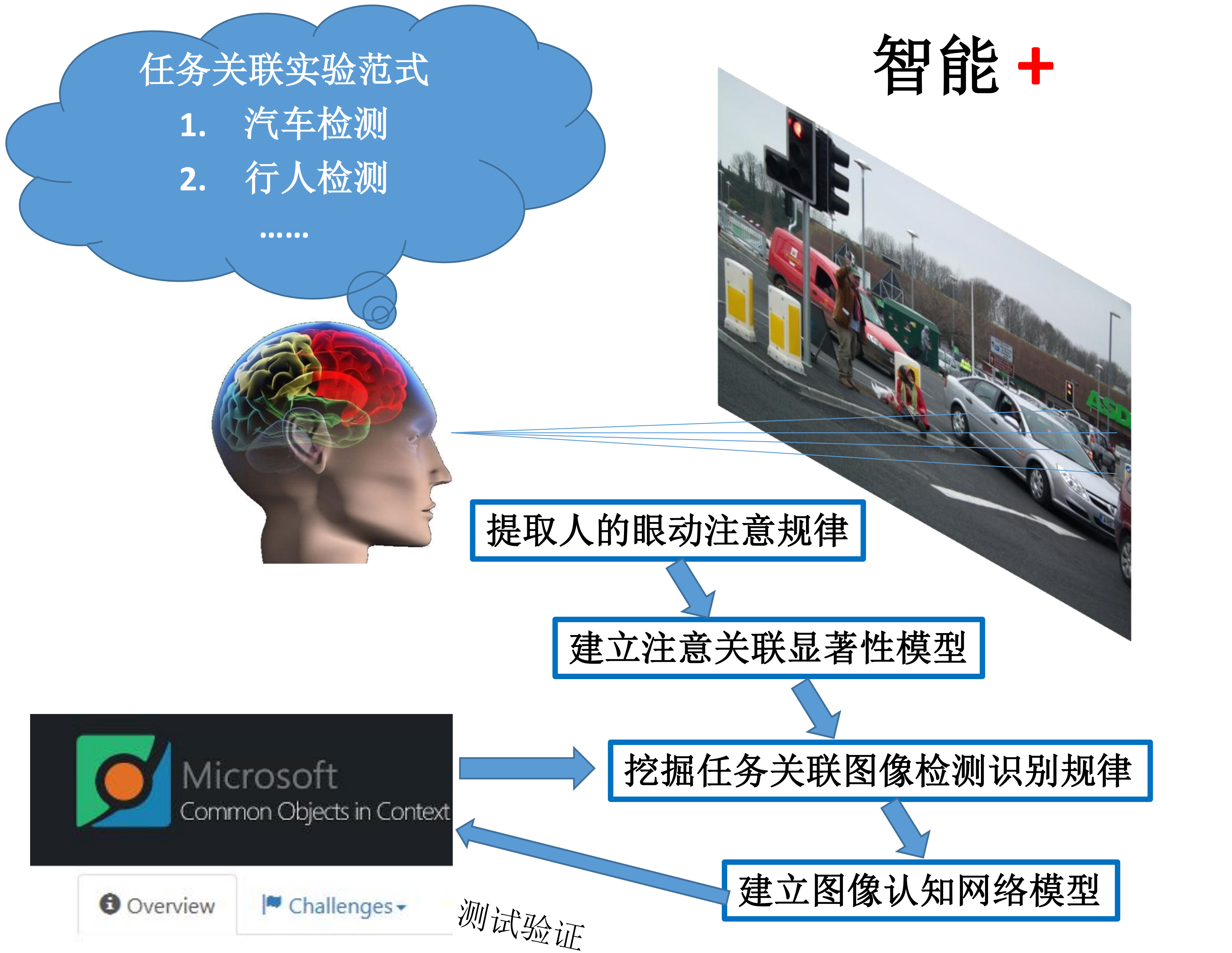

recall < **75%**

COCO:

- 80 classes, ~200K images


recall < **40%** 微软152层**50%**

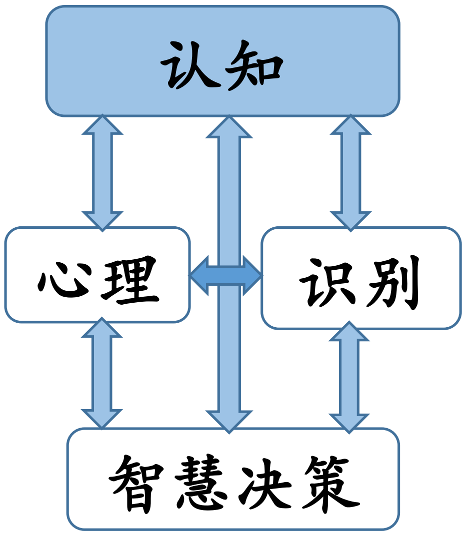
smaller objects, heavy occlusion, high resolution



智能+
任务关联实验范式
1. 汽车检测
2. 行人检测
……
提取人的眼动注意规律
建立注意关联显著性模型
挖掘任务关联图像检测识别规律
建立图像认知网络模型
Microsoft
Common Objects in Context
Overview
Challenges
测试验证

http://3dimage.ee.tsinghua.edu.cn

智能 +



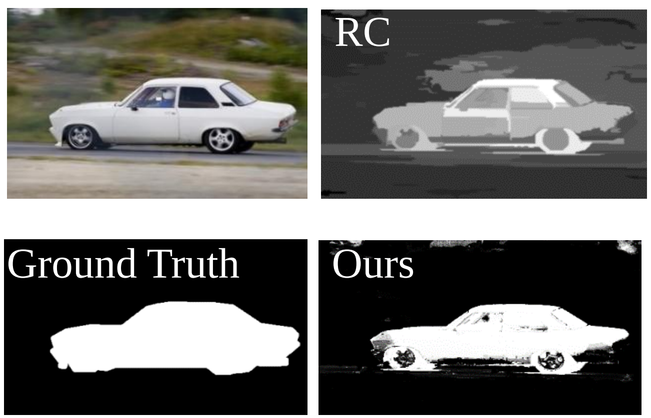
- 让机器学习人的思考模式
  - **显著性物体检测**：语义注意认知模型
  - **部件与结构认知模型**：抵抗遮挡能力
  - **3D 场景物体识别**：适应复杂环境
  - **图像认知心理学**：心理特征量化提取

- 代表性文章与专利

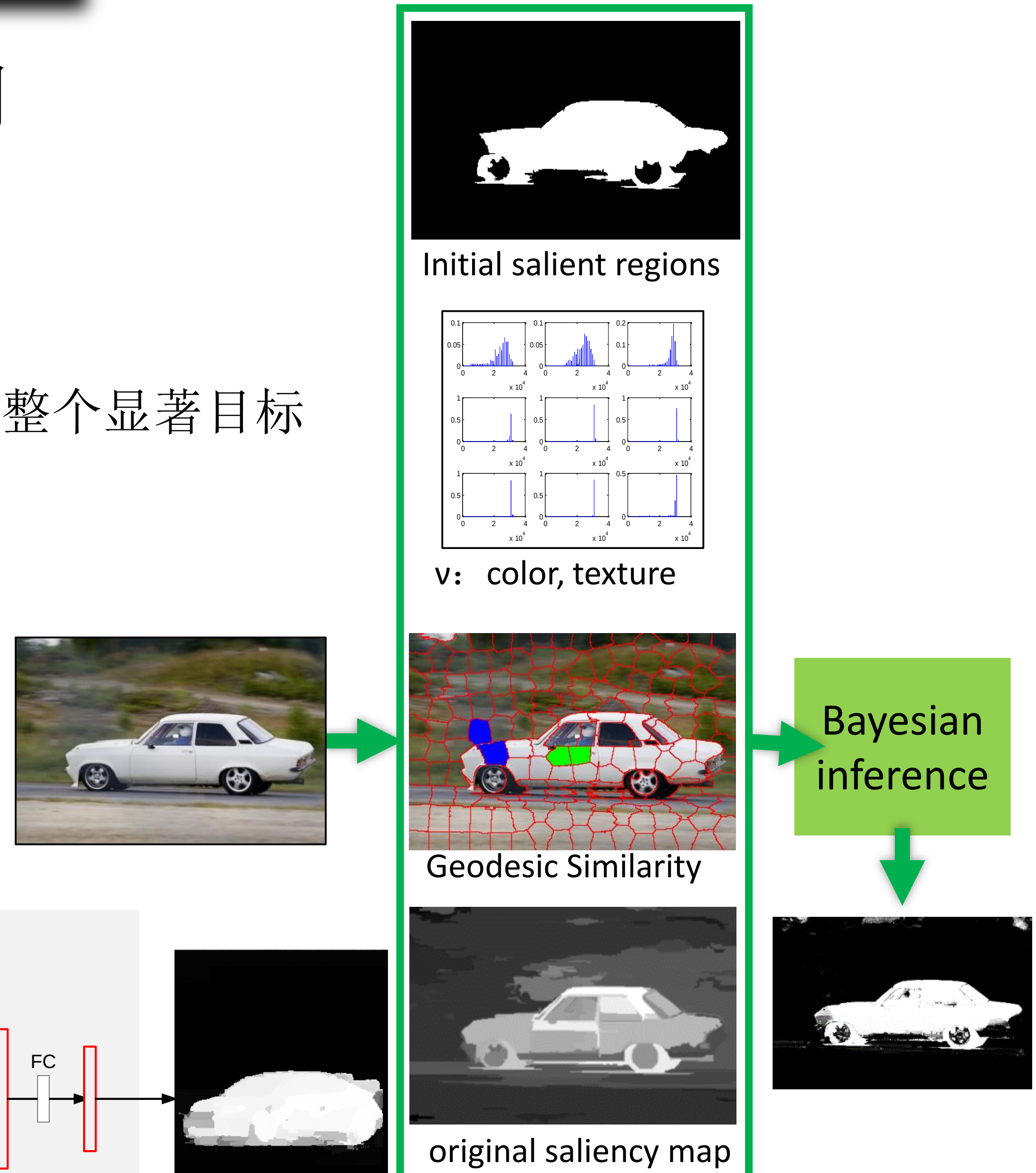
PRL ： Geodesic Weighted Bayesian Model for Saliency Optimization， 2016
PR ： Manifold topological multi-resolution analysis method, 2011
CVPR: Monocular 3D Object Detection for Autonomous Driving, 2016
CVPR: Improving Object Proposals with Multi- Thresholding Straddling Expansion, 2015
NIPS : 3D Object Proposals for Accurate Object Class Detection, 2015
专利： 201310221563.9 一种图像认知心理分析系统
专利： 201310222323.0 基于MMPI心理量表的图像库及其构建方法

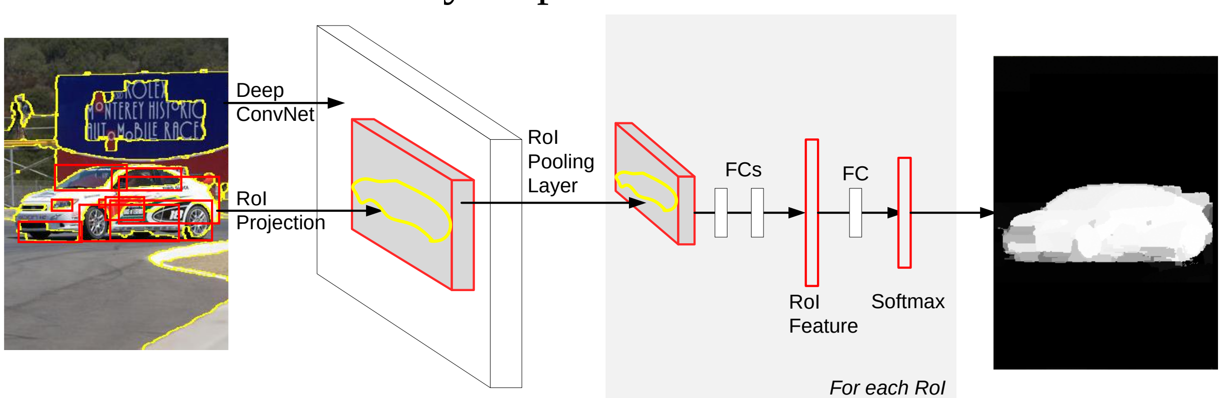
（1）显著性物体检测

- **提出自顶向下的认知模型**
  - 利用贝叶斯框架抑制噪声
  - 利用区域空间关系统一标出整个显著目标



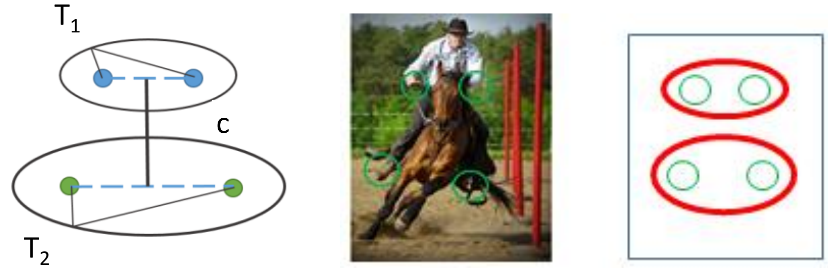
图像的Saliency map：





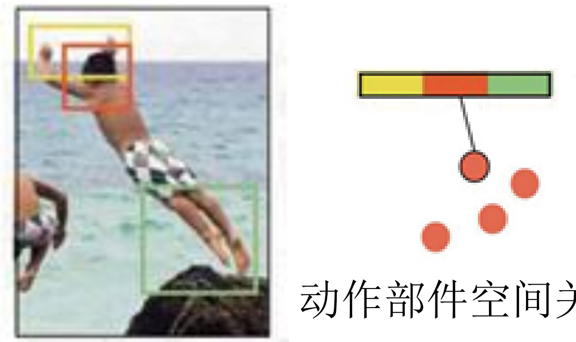
（1）显著性物体检测

（2）部件与结构认知模型



对称对结构模型

- **提出通用对称对结构模型**
  - 建立了动作图象中的对称对结构模型
  - 降低了结构对部件准确检测的要求

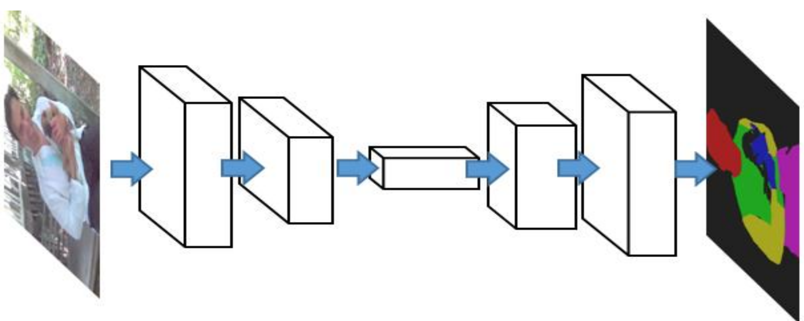


动作部件空间关系

- **提出部件的Top-down空间结构模型**
  - 确定部件的空间关系提高动作检测准确性
  - 建立空间结构神经网络和分级分辨神经网络

Human-level concept learning

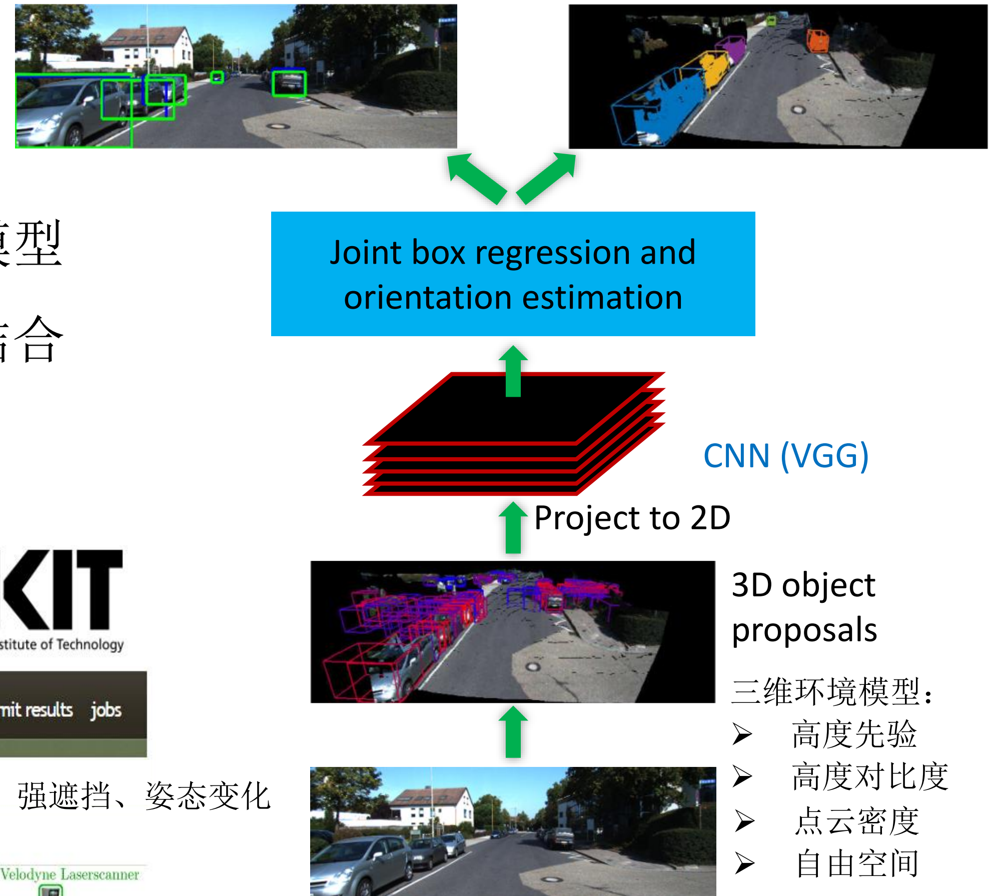
2015.12.12 Science



用全卷积网络确定并定位部件

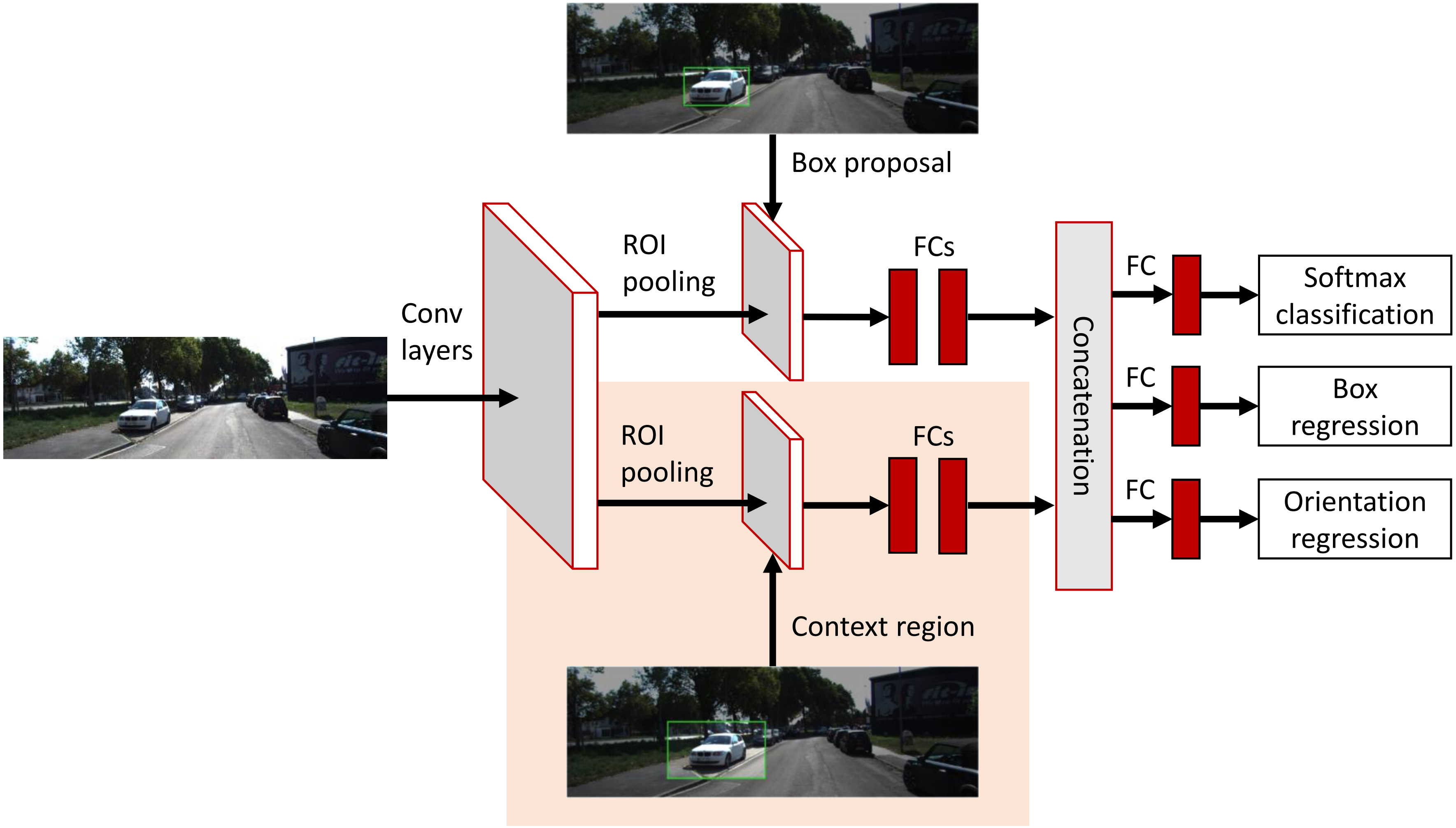
（3）3D 场景物体识别

- 提出三维环境认知学习模型
  - ➢ 建立人在物体检测中的认知模型
  - ➢ 将三维环境先验与深度学习结合
  - ➢ 提出3D物体检测、定位方法



3D Proposal　　Free Space　　Height prior

3D Object Detection Network

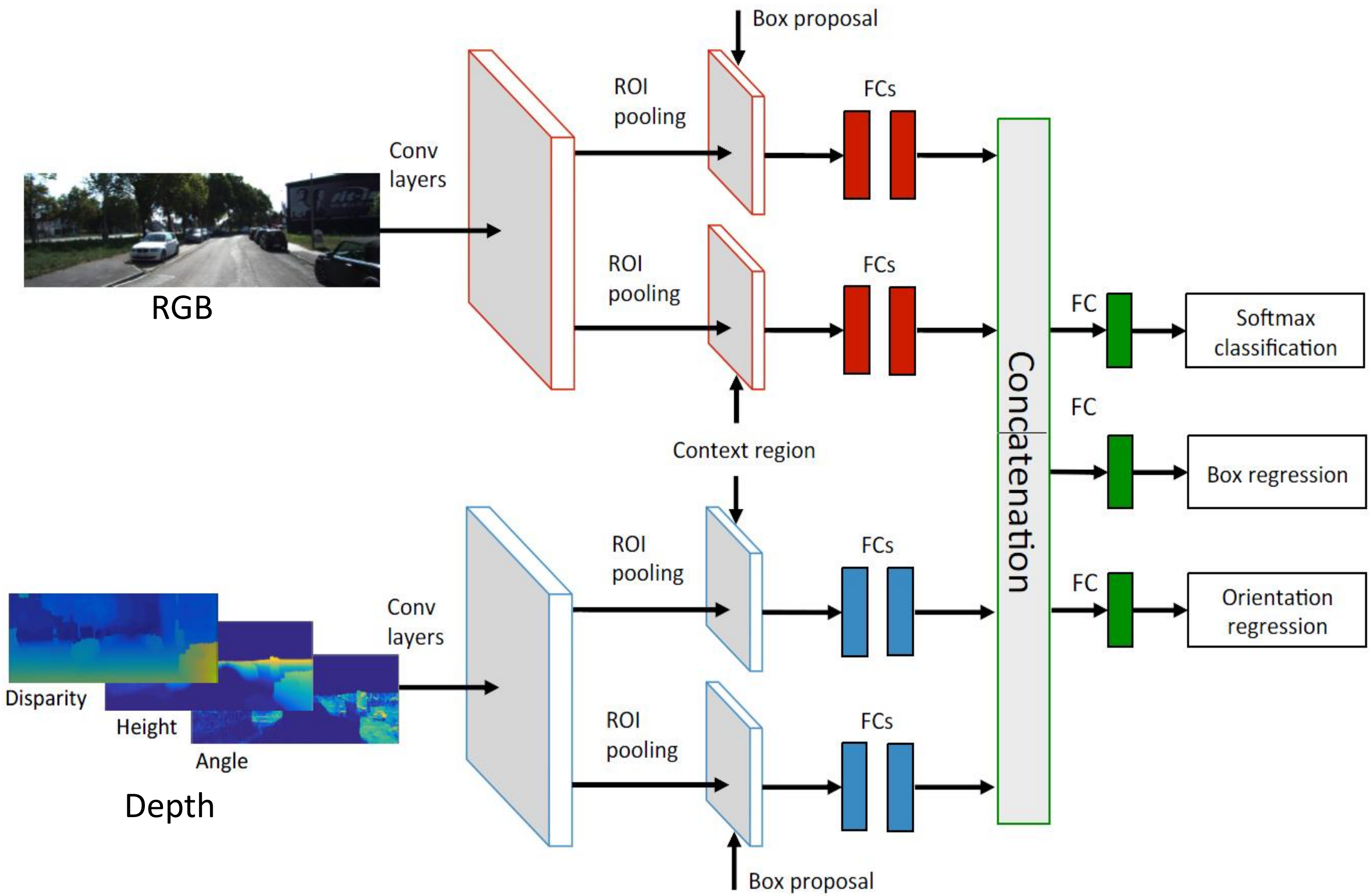


[Chen & Ma, et al. NIPS'15 & CVPR'16]

3D Object Detection Network

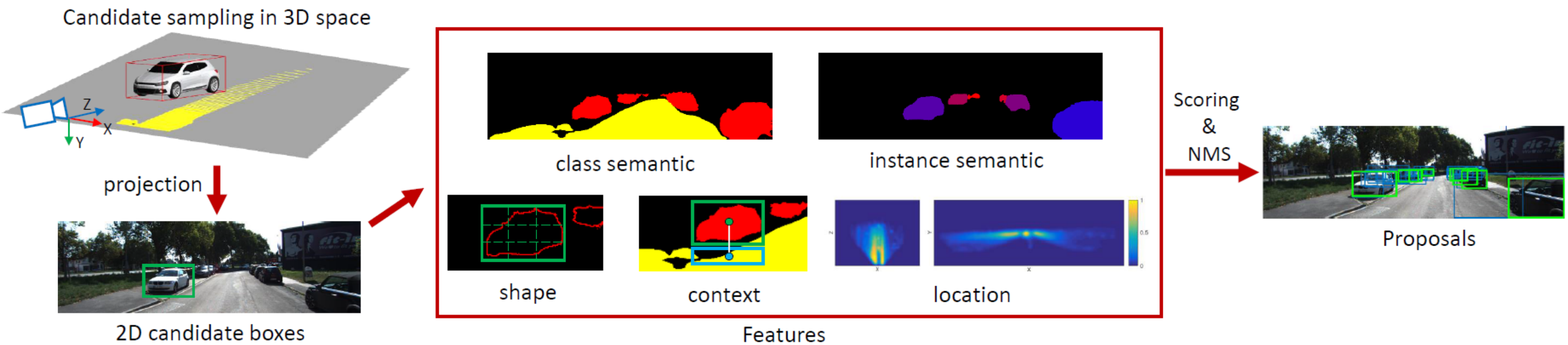
[Chen & Ma, et al. NIPS'15 & CVPR'16]

- 包含上下文信息：Incorporating context information
- 多任务监督：Multi-task supervision
- 多通道特征学习：Multi-stream feature learning



单目三维类目标检测

Monocular 3D Object Proposals - Overview

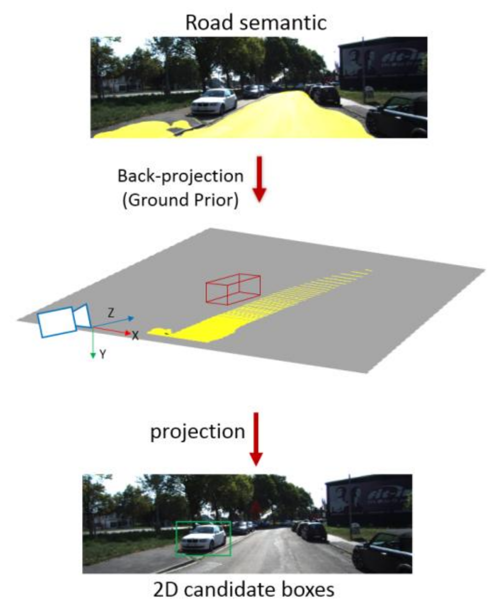


立体/Stereo

- 3D Sampling
- Road Estimation from 3D
- Point Cloud Features
- Exhaustive Search
- Structured SVM

单目/Monocular

- 3D Sampling
- Road Estimation from 2D
- Semantic Features
- Exhaustive Search
- Structured SVM



Results: Object Detection and Orientation Estimation

- **Object detection (AP)**

Car

	Method	Setting	Code	Moderate	Easy	Hard
1	DenseBox2			**89.32 %**	**93.94 %**	**79.81 %**
2	DJML			88.79 %	91.31 %	77.73 %
3	3DOP			88.64 %	93.04 %	79.10 %
	X. Chen, K. Kundu, Y. Zhu, A. Berneshawi, H. Ma, S. Fidler and R. Urtasun: 3D Object Proposals for					
4	SubCNN			88.55 %	90.74 %	77.95 %
	Anonymous submission					
5	CVPR #1408			88.09 %	90.11 %	78.32 %

Pedestrian

	Method	Setting	Code	Moderate	Easy	Hard
1	3DOP			**67.47 %**	**81.78 %**	**64.70 %**
	X. Chen, K. Kundu, Y. Zhu, A. Berneshawi, H. Ma, S. Fidler and R. Urtasun: 3D Object Proposals for A					
2	CVPR #1408			66.34 %	80.25 %	63.41 %
	Anonymous submission					
3	SubCNN			66.13 %	79.13 %	61.27 %

Cyclist

	Method	Setting	Code	Moderate	Easy	Hard
1	3DOP			**68.94 %**	**78.39 %**	**61.37 %**
	X. Chen, K. Kundu, Y. Zhu, A. Berneshawi, H. Ma, S. Fidler and R. Urtasun: 3D Object Proposals for Ac					
2	CVPR #1408			67.03 %	77.06 %	59.87 %
	Anonymous submission					
3	SubCNN			61.98 %	74.40 %	54.75 %

- **Object detection and orientation estimation (AOS)**

Car

	Method	Setting	Code	Moderate	Easy	Hard
1	SubCNN			**87.88 %**	90.49 %	**77.10 %**
	Anonymous submission					
2	DJML			87.51 %	90.67 %	76.33 %
3	3DOP			86.10 %	**91.44 %**	76.52 %
	X. Chen, K. Kundu, Y. Zhu, A. Berneshawi, H. Ma, S. Fidler and R. Urtasun: 3D Object P					
4	CVPR #1408			85.66 %	88.31 %	75.89 %

Pedestrian

	Method	Setting	Code	Moderate	Easy	Hard
1	3DOP			**59.80 %**	**72.94 %**	**57.03 %**
	X. Chen, K. Kundu, Y. Zhu, A. Berneshawi, H. Ma, S. Fidler and R. Urtasun: 3D Objec					
2	SubCNN			59.40 %	72.61 %	54.78 %
	Anonymous submission					
3	CVPR #1408			57.86 %	71.13 %	54.94 %

Cyclist

	Method	Setting	Code	Moderate	Easy	Hard
1	3DOP			**58.68 %**	**70.13 %**	**52.35 %**
	X. Chen, K. Kundu, Y. Zhu, A. Berneshawi, H. Ma, S. Fidler and R. Urtasun: 3D Objec					
2	CVPR #1408			55.73 %	66.64 %	49.75 %
	Anonymous submission					
3	SubCNN			52.06 %	63.74 %	45.93 %

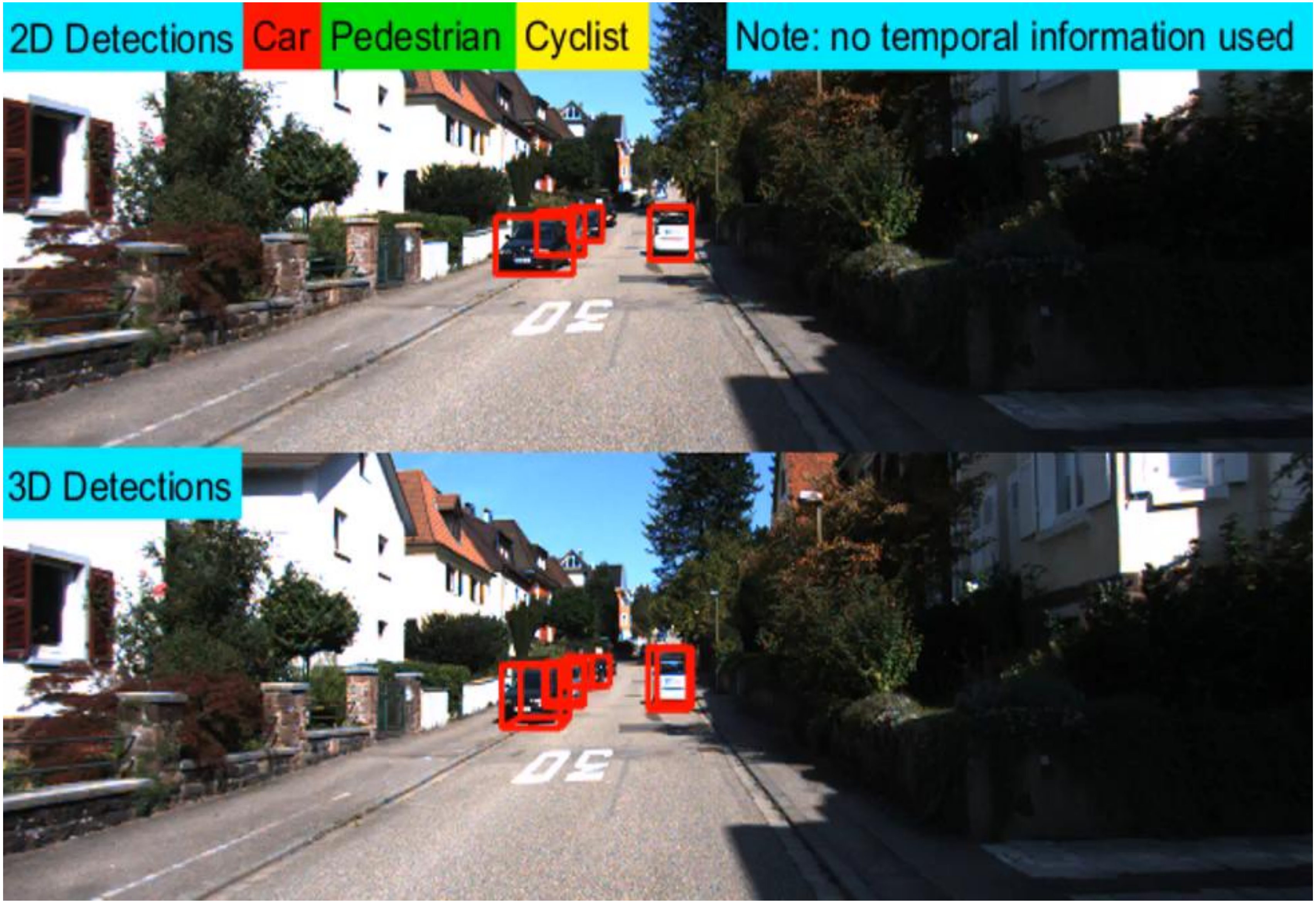
http://www.cvlibs.net/datasets/kitti/eval_object.php

Car	Object detection (AP)			Object detection and orientation estimation (AOS)		
Method	**Easy**	**Moderate**	**Hard**	**Easy**	**Moderate**	**Hard**
SubCat [1]	84.14	75.46	59.71	74.42	83.41	58.83
3DVP [2]	87.46	75.77	65.38	86.92	74.59	64.11
AOG [3]	84.80	75.94	60.70	-	-	-
Regionlets [4]	84.75	76.45	59.70	-	-	-
spLBP [5]	87.19	77.40	60.60	-	-	-
Faster R-CNN [6]	86.71	81.84	71.12	-	-	-
3DOP (Ours)	**93.04**	88.64	**79.10**	**91.44**	86.10	76.52
Mono3D (Ours)	92.33	**88.66**	78.96	91.01	**86.62**	**76.84**

Pedestrian	Object detection (AP)			Object detection and orientation estimation (AOS)		
Method	**Easy**	**Moderate**	**Hard**	**Easy**	**Moderate**	**Hard**
DPM-VOC+VP [1]	59.48	44.86	40.37	53.55	39.83	35.73
FilteredICF [2]	67.65	56.75	51.12	-	-	-
DeepParts [3]	70.49	58.67	52.78	-	-	-
CompACT-Deep [4]	70.69	58.74	52.71	-	-	-
Regionlets [5]	73.14	61.15	55.21	-	-	-
Faster R-CNN [6]	78.86	65.90	61.18	-	-	-
Mono3D (Ours)	80.35	66.68	63.44	71.15	58.15	54.94
3DOP (Ours)	**81.78**	**67.47**	**64.70**	**72.94**	**59.80**	**57.03**

Cyclist	Object detection (AP)			Object detection and orientation estimation (AOS)		
Method	**Easy**	**Moderate**	**Hard**	**Easy**	**Moderate**	**Hard**
DPM-VOC+VP [1]	42.43	31.08	28.23	30.52	23.17	21.58
Mono3D (Ours)	76.04	66.36	58.87	65.56	54.97	48.77
3DOP (Ours)	**78.39**	**68.94**	**61.37**	**70.13**	**58.68**	**52.35**

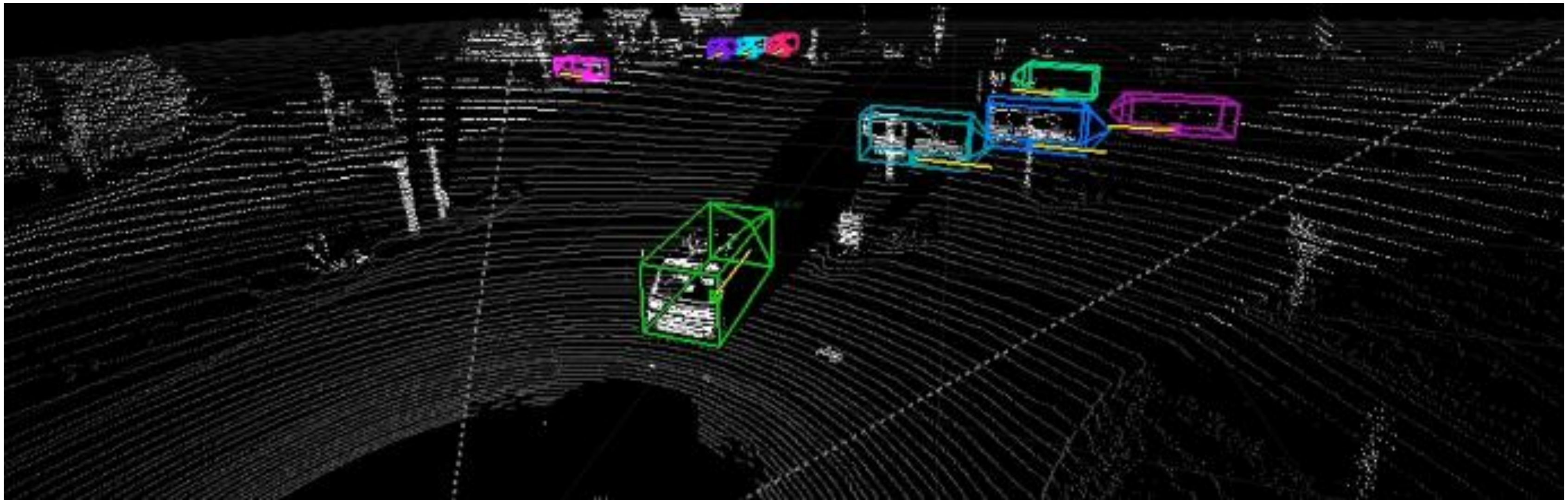
复杂干扰环境中的三维目标识别

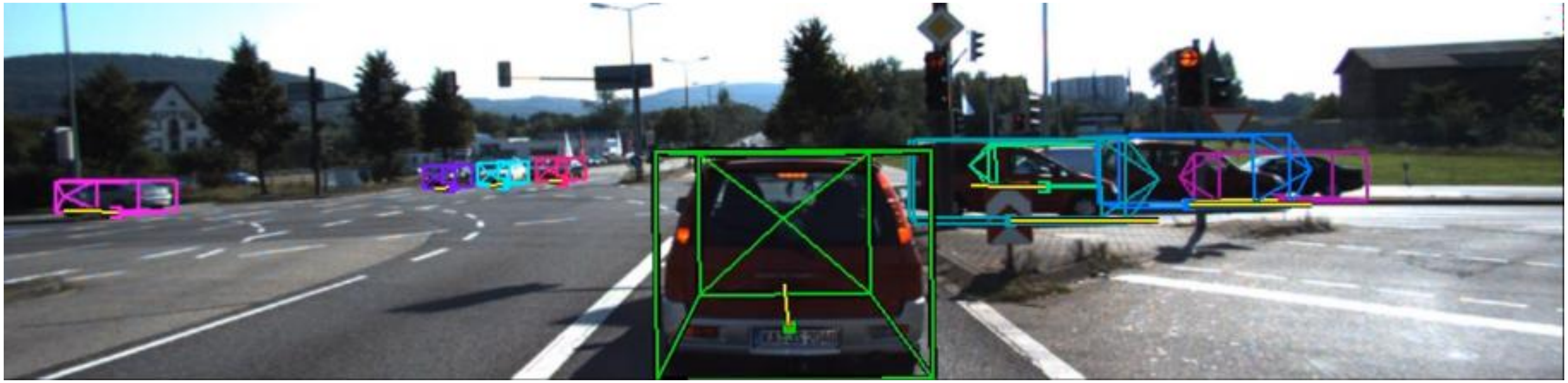


智能 ++ 高级认知：多模型&多任务&多视角

Multi-Modal & Multi-Task & Multi-View

- **Multi-Modal**





高级认知：多模型&多任务&多视角

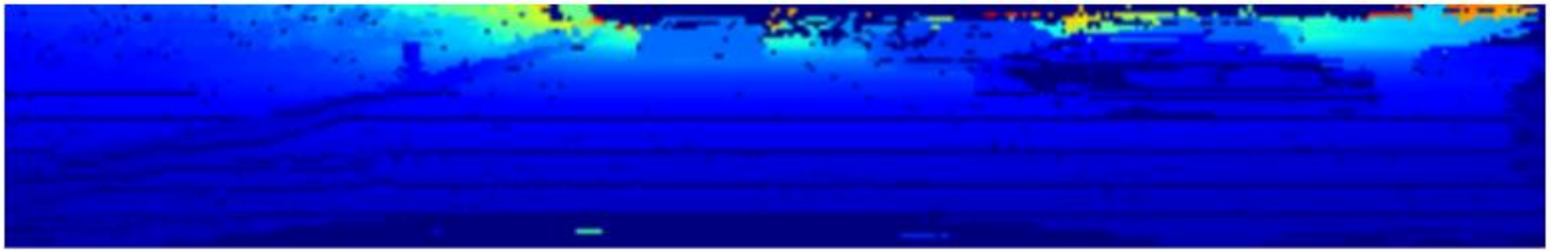
Multi-Modal & Multi-Task & Multi-View

- **Multi-Task**

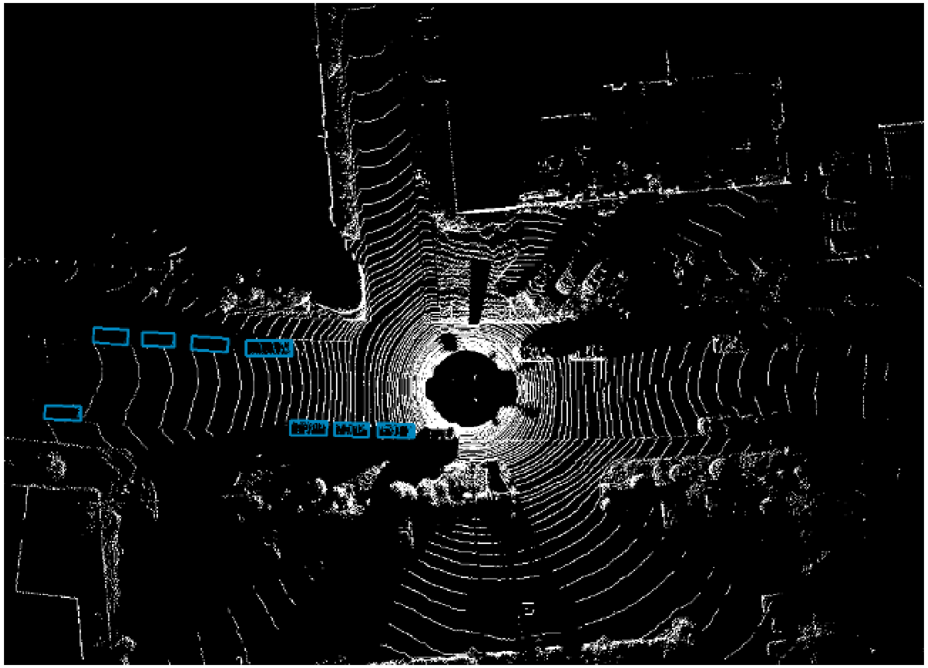
高级认知：多模型&多任务&多视角

Front view

- **LIDAR**：Input: depth, height, distance
- **RGB**



Bird's eye view

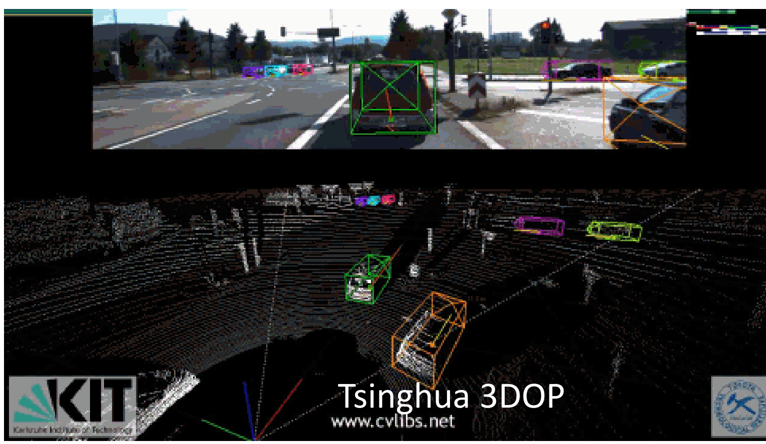


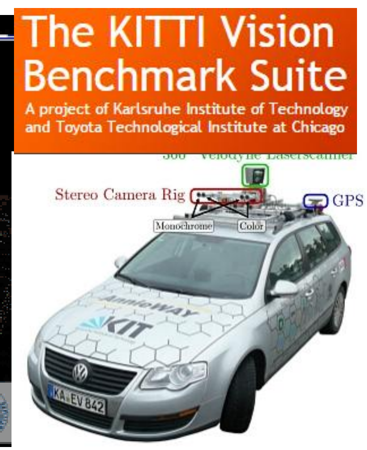
- Input representation
- Sparse points
- Extra small objects
- Oriented bounding boxes

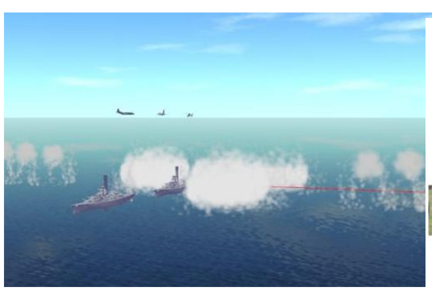
应用：智能驾驶与视觉感知

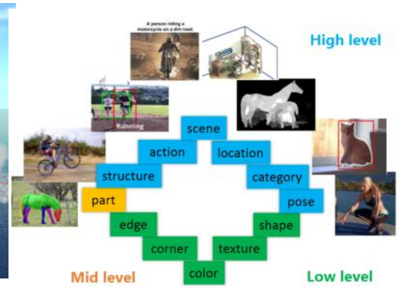
➤ 小目标 ➤ 强遮挡 ➤ 任意姿态 ➤ 人

- 建立了**国际上唯一的**图像认知心理评测系统
- 2015自动驾驶KITTI国际评测，共六项获**四项世界第一**
- 复杂环境目标检测识别技术服务于**航空、航天、交通**







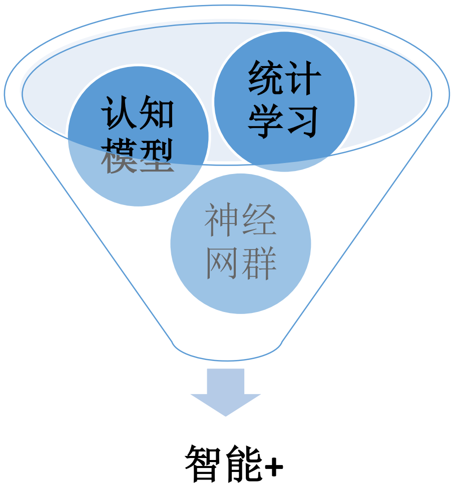


➤ **国家自然基金：**基于图像认知的心理特征提取与分析

➤ **专项重点基金：**适应复杂环境的成像导航

➤ **国家重点研发计划：**智能电动汽车感知、决策与控制

智能+



发表的相关文章：

http://3dimage.ee.Tsinghua.edu.cn

ICIP 2014: Learning a compact latent reprecentation of the bag-of –parts model
ICCVG 2014: Protected Pooling Method Of Sparse Coding In Visual Classication
ICIP 2015: Geodesic weighted Bayesian model for salient Object Detection.
CVPR 2015: Improving Object Proposals with Multi-Thresholding Straddling Expansion
NIPS 2015: 3D Object Proposals for Accurate Object Class Detection
CVPR 2016: Monocular 3D Object Detection for Autonomous Driving
ICIP 2016: Object Detection Via Fast R-CNN And Low-level Cues
ICIP 2016: Region Candidate Combination for Action Recognition.
PRL 2016: Geodesic Weighted Bayesian Model for Saliency Optimization
PRL 2016: Semantic parts based top-down pyramid for action recognition
PR 2016: Generalized Symmetric Pair Model for Image Recognition, Pattern Recognition (修改中)

Thank You

http://3dimage.ee.Tsinghua.edu.cn

电子工程系
Electronic Engineering
清華大學
Tsinghua University