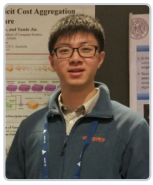
于立冬

151-6523-1608 | yvlidong@gmail.com | 北京
Google Scholar | lidongyv.github.io



教育经历

北京理工大学（985） 2016.08 — 2019.06
计算机科学与技术 · 硕士 · 计算机学院 北京

中国海洋大学（985） 2012.08 — 2016.06
计算机科学与技术 · 本科 · 计算机学院 青岛

工作经历

新石器慧通（北京）科技有限公司 2025.08 — 至今
AD 算法负责人 · AD 二部 北京
算法团队 38 人，~8,700 卡

- **Perception**：Group query-based 多任务模型，统一 object detection、lane detection 与 occupancy network；基于 E2E 的红绿灯检测与导航模型
- **PNC 与两阶段 Planner**：静态轨迹预测 + 后置动态轨迹优化
- **E2E 与世界模型**：两阶段 pretrain（LLM 蒸馏 + 感知模型 inception）；基于 AR 的轨迹预测，多任务 FT 与 CoT debugger；基于 RL 的 post-training
- **Infra**：数据闭环；离线大模型自动标注；多任务训练自动配平；多平台存储与训练；闭环评测与仿真器

NVIDIA（英伟达） 2024.07 — 2025.08
Principal Engineer · 自动驾驶部门 Free IC 北京

- **模型**：Roadnet（16 人，~1,200 卡，3 个月）HD→Mapless，segment→query→token-based；Planner（13 人，~1,500 卡，3 个月）经典 PNC→model-based，two-stage E2E with AR；E2E（13 人，~3,500 卡，6 个月）Alpamayo 量产 ORIN 交付奔驰 SOP，one-stage E2E with Cosmos CoT 与 Alpamayo WM-RL
- **Infra**：感知持续训练（数据自动配平 + 挖掘框架）、蒸馏驱动的多任务训练优化；E2E-Alpamayo 的 Backbone 预训练 / FT / post-training 框架

北京地平线机器人技术研发有限公司 2023.06 — 2024.07
静态感知算法负责人 · 土星五号项目 北京
负责 15 人模型 + 5 人后处理团队，~1,200 卡

- 负责 static perception 的 model + post-process + data infra；lane detection 由 segment-based → query-based → group-wise multi-task；static E2E + navi model
- **Infra**（开发 7 人 + 外包 5 人 + 下游 20 人）：0→1 搭建数据闭环；重构架构（issue 解决效率 D+10→D+3）；自动化标注工具（4D GT），1k→3w 路口/月；data auto mining，冗余 80%→5%

深圳元戎启行科技有限公司 2022.04 — 2023.06
静态感知模型负责人 北京
核心团队 4 人，600 卡

- 2023 年完成国内第一套 Mapless 无图方案，交付问界 POC；Lane segment 方案支持城区点到点与高速自动巡航，匝道/路口通过率 >85%；lane segment → query-based → AR-based lane detection
- **Infra**：>2 亿 img 数据挖掘平台，~1,200w frame 自动挖掘；多任务训练 auto-balance 框架

上海蔚来汽车有限公司 2021.03 — 2022.04
主任工程师 · 自动驾驶算法 北京

- 国内第一代 AD 自研算法（NIO ET7、ES8）；Data-Driven 的 Mono3D detection-tracking 模型；object detection / lane detection 的数据闭环与模型评测框架

驭势科技（北京）有限公司 2019.05 — 2020.04
高级研究员 · 南京 AI 研究院 南京

- 2019 年国内第一家实现 Shadow Mode 数据闭环系统 + IL-based E2E 上车；面向 AD 的 imitation learning、RL、数据生成与仿真；3D 目标检测与跟踪联合

发表论文

1. 3D Visual Illusion Depth Estimation. **NeurIPS 2025**
2. Diving into the Fusion of Monocular Priors for Generalized Stereo Matching. **ICCV 2025**
3. Sparse Point Guided 3D Lane Detection. **ICCV 2023**
4. Parameterized Cost Volume for Stereo Matching. **ICCV 2023**
5. FoggyStereo: Stereo Matching with Fog Volume Representation. **CVPR 2022**
6. Deep Stereo Matching with Explicit Cost Aggregation Sub-Architecture. **AAAI 2018**
7. Content-Aware Inter-Scale Cost Aggregation for Stereo Matching. **arXiv 2020**
8. DRL-based Trajectory Tracking for Motion-related Modules in Autonomous Driving. **arXiv 2023**

专利

1. 三维重建方法、装置、系统和存储介质 （CN201980002779.1）

2. 图像预测及车辆行为规划方法、装置和系统及存储介质 （CN202010196263.X）
3. 多目标跟踪方法、装置和系统及存储介质 （CN201911334843.4）
4. 单目标跟踪方法、装置和系统及存储介质 （CN201911194289.4）
5. 生成对抗网络的训练方法和系统、电子设备和存储介质 （CN201980002600.2）

项目详情

Neolix —— AD 算法

2025.08 — 至今

- **Perception** —— 重构单任务 / 多任务 task formulation
  - 使用 Group query-based model 统一 single-task / multi-task 训练，从 multi-head 设计转为数据友好的 single-head 多任务设计
  - One query-based Model 覆盖 object detection、lane detection、occupancy network、E2E 红绿灯检测与导航模型
  - single head 可自设 data balance 而不影响其他任务；按 BEV token / img token 在 token 级别（而非 task-loss 级别）rebalance task weight
  - 重构感知任务以适配 E2E：Lidar 前融合、object detection、MOT 升级为 object prediction，future tokenizer 支持 pre/post-training
  - lane detection、红绿灯检测升级为 navi-based static E2E model，降低网络学习难度，为前置决策提供 strong prior
- **E2E 与世界模型（WM）**
  - 两阶段 pretrain —— 从 Cosmos / VGGT 进行 LLM 蒸馏（CoT 能力、3D 几何估计、3D OD / lane 任务、failsafe 与场景识别）；perception model inception（提供 AD task-force 知识，服务 RL 与 vis debugger）
  - 可控的迭代训练发版 —— frozen head 预训练 backbone；frozen backbone 微调 perception head；支持随时加数据
  - 基于自回归（AR）的轨迹预测 —— pre/post perception/perceiver 的多任务 FT；CoT debugger；多模态 diffusion；RL post-training
  - 通用世界模型 —— pretrain（backbone LLM 蒸馏，Cosmos→CoT 的 language inception；D4RT 提供统一并行 AR 与 3D insight）；render（面向 generalized generation 而非单场景重建，DiT 与 3D-GS 仅用于 render）；conditional inception（language / layout / traj）
- **Infra** —— AI native 与 RaaS 转型
  - VLM 自动 triage；agent skills 驱动评测与仿真；个人 AI agent 提效会议；Vibe coding 驱动的自动研究
  - E2E 与 WM 的数据闭环：离线感知模型自动标注；VLM 自动挖掘
  - agent 驱动的前后端开发与人工客服；仿真器 auto checker

NVIDIA —— AD / E2E / 世界模型

2024.06 — 2025.07

- **Perception Model → 两阶段 E2E**
  - segment-based lane detection：所见即所得，优化超分、提高几何精度、BEV 多尺度融合、动态多分辨率 DCN
  - query-based lane detection（静态 E2E）：摆脱"所见即所得"formulation，增强脑补能力，输入 navi info，新增过路口、split/merge、lane change 等 navi model 前置决策
  - token-based lane detection：协助两阶段 E2E 升级至 one-stage Alpamayo，以 AR 方式 tokenize 所有 perception 元素，支持前置 conditional SA 与后置 debugger vis
- **PNC → one-stage Alpamayo**
  - planning 模块模型化；single-mode → multi-mode regression；regression → AR → diffusion formulation
  - perception 后置、one-stage 升级：Cosmos / Qwen / Llama backbone 蒸馏；CoT、3D geo、AD 多任务 inception 训练；去除 perception tokenizer（后置为 debugger 任务）；planning 由 based-on-perception 改为 based-on-image-stream
  - planning FT 与 post-training：AR / diffusion / 多模态 traj prediction；RL post-training；与 Cosmos-WM 的闭环训练
- **世界模型** —— Cosmos pretrain → render 研究；从 VGGT、Qwen 蒸馏

地平线 · 土星五号 —— 道路结构认知

2023.08 — 2024.08

轻地图 PNC 方案与众包建图下的道路认知模块。

- **算法研发 —— 连续要素（车道线）**
  - 路沿采用"所见即所得"策略，模型仅输出车道线与路沿（不输出中心线），针对路口遮挡等场景设计数据清洗与标注规则
  - 以 query-based 方案替换 segment-based，解放后处理、人力前移，计算效率持平，F1 提升 15%
  - 路口两侧横向及远处弱可见区域 recall 大幅提升；相比 entry-point 策略可直接构建车道
  - query-based 时序策略将 tracking 前移至网络部分；以 PV refinement 降低对 BEV feature 的强依赖并引入新多模态融合方式；以 query 分组支持 case 级别网络精调
- **离散要素（arrow / crosswalk / junction）**
  - 以 query-based 方案替换二阶段检测，点数持平，计算效率提升 30%+
  - 以 query 分组拉齐单任务 / 多任务指标差距，支持同源数据的 one-model 多任务训练与联调
- **数据闭环**
  - 构建道路结构全要素、多任务、同源时序数据集；针对 query-based 方案优化数据采集、标注规则与数据配平
  - 联调开发自动化标注工具（4D GT）；参与自动化数据采集系统、路测链路系统、车云协同系统、道路采集覆盖系统等数据大闭环系统的 0—1 架构与实现

元戎 —— MapFree 下的车道线检测

2022.05 — 2023.06

Data-Driven 的车道线检测系统。

- **GT 生成**：搭建 HDmap → Local BEV 的 GT 生成系统（200 万 image，1,200 万 lane instance）
- **投影 GT 友好的检测模型**：优化 segment-based 车道线检测（AP 13%→60%，EPE 1.5m→0.18m）；标注规则定义与数据清洗；direction head 连线，SR 提高精度
- **query-based 车道线检测**：基于"车道-车道线-点链"的多级回归算法；数据拓扑定义、算法开发、评测、可视化工具全线开发
- **基于 language model 的拓扑学习**：自回归模型关键点生成；数据拓扑定义与算法开发
- **segment-based 拓扑学习**：基于 MaskFormer 的拓扑关键点分割；"逆转 transformer"将 query-based 点链生成转化为 segment 算法
- 数据驱动的自动模型迭代与评测

NIO —— 3D OD 与车道线量产

2021.03 — 2022.04

- **工作内容**
  - 关联后处理算法（rule-based）开发，支持 7 cam + 1 lidar 跟踪关联
  - 关联学习化：设计 learnable cost metric 与 matching scheme 替代规则与二分图匹配
  - ReID 数据标注与 pseudo label 生成
  - 后处理算法车端部署测试，交付 ET7 NAD & SUD 功能
  - Monocular Depth Estimation 数据标注、算法开发、量化部署评测
  - Ground Model 数据标注、算法开发、量化部署、路测回测
  - 3D 车道线方案支持与后处理研发
- **结果**
  - 在 20% 3D 错误率下提供关联算法：0.03 ID switch、86% Recall，整体性能超过 EQ4
  - 150m 外远距离小物体关联分裂率 0.1%；拥堵、停车场景下的稳定关联
  - 全图 dense depth 6.8% rel error，地面 1.3% rel error
  - 支持量产车通过 HOMO 测试，LKA、LKW 功能测试通过
  - 支持 ET7 SUD & NAD 功能路测与量产部署

研究项目（Research Projects）

2021.03 — 2025.08

- Deep stereo matching with explicit cost aggregation ↗
- FoggyStereo: stereo matching with fog volume representation ↗
- Diving into the fusion of monocular priors for generalized stereo matching ↗
- 3D visual illusion depth estimation ↗
- Sparse point guided 3D lane detection ↗

UISEE —— AD 研究

2019.07 — 2020.12

- Occluded 3D object detection and tracking with RepPoint formulation ↗
- Clustering-based feature disentanglement for ReID, SOT, and MOT ↗
- Relative navigation and shadow-mode data collection with trajectory imitation ↗