Zhihui Wang

Profile

Date of Birth: 09/12/1994

Work Experience: 6 Years

Research Direction: Computer Vision, Deep Learning

Job Seeking Status: Employed

Contact Information: (+1)832-875-9593

E-mail: jillian_wzh@163.com

Work experience

DiDi Senior Algorithm Engineer Map Division-Traffic Big Data-Visual Computing Group 07/2019 - present

Educational Background

09/2016-06/2019 Master Object tracking, object detection Dalian University of Technology

Advisor: Professor Wang Dong and Professor Lu Huchuan

CCS: Online Single Person Tracking for Unmanned Aerial Vehicles: Benchmark and New Baseline

09/2012-06/2016 Bachelor Electrical and Computer Engineering Dalian University of Technology

Patents and Publications

- Decoupling with Entropy-based Equalization for Semi-Supervised Semantic Segmentation **IJCAI-2023**
- 2nd Place Solution for Waymo Open Dataset Challenge-Real-time 2D Object Detection **CVPR WS-2020**
- Robust and Fast Vehicle Turn-counts at Intersections via an Integrated Solution from Detection, Tracking and Trajectory Modeling **CVPR WS-2019**
- Online Single Person Tracking for Unmanned Aerial Vehicles: Benchmark and New Baseline **ICASSP-2019**
- Online Vehicle Tracking in Aerial Imagery **IScIDE-2017**
- Traffic accident recognition method, device, electronic device and medium **ID: CN112926575A**
- Vehicle counting method and system, data processing equipment and intelligent shooting equipment **ID: CN111652912B**
- Method and apparatus for presenting road information **ID: WO2022156553A1**
- Method and device for detecting bus lane, electronic equipment and storage medium **ID: CN112733793A**

Awards

2022 Beijing Middle Surveyor

2021 Didi Technology Co., Ltd. - Map and Public Transportation Team - Polaris Star Award

2019 Didi Technology Co., Ltd. - Map and Public Transportation Team - Shining Star Award

2016-2018 **National Level Graduate Fellowship**, **National Level Graduate Scholarship**

2016 Dalian Lingshui Scholarship, Outstanding Graduate of Dalian University of Technology

2012-2016 Dalian University of Technology Study Scholarship

Project Experience

- **Didi Clairvoyance-Road Dynamic Event Mining**
  - **End-to-End System Development:** Built a terminal + cloud-based dynamic road event detection system from the scratch, leveraging Didi's vast repository of dashcam footage. Led the design and execution of the full-stack technical solution.
  - **Edge-Device Optimization:** Overcame on-device computational limitations by deploying lightweight neural networks (ShuffleNet V2, MobileNet V3) for real-time feature recognition. Enhanced model efficiency through knowledge distillation, neural architecture search (NAS), and other compression techniques.
  - **Cloud-Based AI Processing:** Developed a high-accuracy server-side event verification system using:

FCOS-based object detection for feature recognition, Swin Transformer-powered scene understanding, Road topology reconstruction via semantic segmentation of road features

- **Data closed loop:** Build a data mining system that combines edge reporting + real-time collection + server verification + model optimization. Through the data closed loop, rapid evaluation and performance optimization of small edge models can be achieved.
- **Business Impact:** Optimized Mobile Inference: Achieved <5% CPU utilization on ARM A53 chips while maintaining 98%+ accuracy in accident detection, enabling fully automated deployment. Metrics: 5,000+ high-priority events processed daily (data quality≥95%); 15,000+ events mapped per day (data quality≥ 90%); Performance metrics consistently outpaced industry rivals in efficiency and reliability.

- **Mobile deployment of the visual odometry AI model Project leader of two members**
  - Assist in building AI models to accelerate the visual odometry in crowdsourcing mapping processes;
  - In-depth analysis Feature point extraction algorithm (Superpoint) and feature matching algorithm (SuperGlue). Based on the algorithm characteristics and time consumption of each module, backbone optimization, Transformer lightweighting, TensorRT quantization, computational graph optimization and other solutions are used to improve model efficiency;
  - **Business Impact:** SuperPoint Inference Time: Reduced from 531 ms to 10.5 ms, a 50x+ speedup; SuperGlue Inference Time: Cut from 1,725 ms to 29 ms, a 59x+ acceleration; Accuracy vs. Efficiency Trade-off: The lightweight model maintains a reconstruction error of < 50 cm, meeting all business requirements while drastically outperforming the original large model in speed.
- **Multimodal scene recognition Project leader of two members**
  - Investigate the application of large visual models in business scenarios. Use multimodal models such as EVA-CLIP-E and Intern VL to mine traffic accident scenes.
  - Business Impact: Achieved 95%+ precision with 90%+ recall rate in accident detection; Implemented fully automated triggering system for incident reporting.
- **HD-Map Boundary Style Recognition Project leader of two members**
  - **System Development:** Pioneered Lane-Style Recognition: Built from scratch (0→1) a lane-boundary intelligence system capable of detecting and updating lane-level map styles (left/right/shoulder/median).
  - **Technical Approach:** Multi-Model Analysis: Combined lane detection and road-feature segmentation models to classify boundary styles (e.g., solid/dashed lines, barriers).
  - **Continuous Learning:** Leveraged millions of daily images to automatically detect style changes and improve update accuracy via secondary validation workflows.
  - **Business Impact:** 30% auto-processing rate for right-side/shoulder styles saves 8+ man-days/month; 40%+ update conversion rate (minimizes manual verification).
- **Other business**
  - **Multi-Sensor Fusion Detection:** Developed a HydraNet-based fusion architecture integrating multi-modal detection capabilities including: traffic sign recognition, vehicle attribute analysis, electronic surveillance detection, traffic light state classification
  - **Regulatory Sign Intelligence System:** led end-to-end development of 100+ prohibition sign detection models for global deployments, driving: continuous model optimization cycles, large-scale traffic sign data mining, production-grade detection pipelines
  - **Privacy-Preserving Data Processing:** established automated image desensitization framework leveraging: targeted object detection for sensitive road features, OCR-based text redaction, geo-sensitive area masking.