



IEEE
COMPUTER
SOCIETY

***3rd Workshop on Vision Datasets Understanding
in conjunction with CVPR 2024***

***Best Paper Honorable Mention Award
presented to***

**DTLLM-VLT: Diverse Text Generation for Visual Language Tracking
Based on LLM**

Xuchen Li¹ Xiaokun Feng^{1,2} Shiyu Hu^{1,2} Meiqi Wu³ Dailing Zhang^{1,2} Jing Zhang¹ Kaiqi Huang^{1,2,4}

¹CRISE, Institute of Automation, Chinese Academy of Sciences

²School of Artificial Intelligence, University of Chinese Academy of Sciences

³School of Computer Science and Technology, University of Chinese Academy of Sciences

⁴CAS Center for Excellence in Brain Science and Intelligence Technology

José Lezama

Qiang Qiu

Liang Zheng