ALVR 2026

Workshop on Advances in Language and Vision Research

Proceedings of the Workshop

July 3, 2026

The ALVR organizers gratefully acknowledge the support from the following sponsors.

Gold

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
317 Sidney Baker St. S
Suite 400 - 134
Kerrville, TX 78028
USA
Tel: +1-855-225-1962
acl@aclweb.org

ISBN 979-8-89176-398-2

Introduction

We are excited to welcome you to the 4th Workshop on Advances in Language and Vision Research (ALVR), co-located with ACL 2026 and held on July 3, 2026, in San Diego, California.

ALVR provides a dedicated forum for research at the intersection of natural language processing and computer vision. This year's program features five invited talks from leading researchers, six spotlight presentations, and twenty-three poster presentations, spanning a wide range of topics at the frontier of language and vision research.

We received 47 submissions this year. After a thorough review process, we accepted 29 papers in total: 23 as archival papers and 6 through our non-archival track, which welcomes work concurrently under submission or recently published elsewhere. The archival acceptance rate is approximately 49%. Each submission received at least three reviews, with area chairs providing guidance to ensure the quality and breadth of the final program. We thank our area chairs—Yue Fan, Syrielle Montariol, Alane Suhr, Xin Eric Wang, and Qianqi Yan—for their careful deliberation, and all program committee members for their thorough and timely reviews.

We are grateful to our five distinguished invited speakers: Raymond J. Mooney (University of Texas at Austin), Lianhui Qin (UC San Diego), Amir Zadeh (Lambda Labs), Mohit Bansal (University of North Carolina at Chapel Hill), and Jiajun Wu (Stanford University), for sharing their insights and inspiring the community.

We gratefully acknowledge the generous support of our sponsor, Lambda Labs (Gold level), whose contribution has been instrumental to the success of the workshop.

We extend our sincere thanks to all authors who submitted to ALVR 2026, and to all attendees for their enthusiasm and participation. It is the community's energy and commitment that makes this workshop a rewarding venue for advancing research at the intersection of language and vision.

Qianqi Yan, Lead Organizer

Syrielle Montariol, Yue Fan, Jing Gu, Jiayi Pan, Manling Li, Parisa Kordjamshidi, Alane Suhr, and Xin Eric Wang, Co-Organizers

Organizing Committee

Organizers

Qianqi Yan, University of California, Santa Barbara, USA
Syrielle Montariol, University of California, Berkeley, USA
Yue Fan, Autodesk Research, USA
Jing Gu, xAI, USA
Jiayi Pan, xAI, USA
Manling Li, Northwestern University, USA
Parisa Kordjamshidi, Michigan State University, USA
Alane Suhr, University of California, Berkeley, USA
Xin Eric Wang, University of California, Santa Barbara, USA

Program Committee

Area Chairs

Qianqi Yan, University of California, Santa Barbara
Yue Fan, Autodesk
Syrielle Montariol, ISIR, Sorbonne Université, France
Alane Suhr, University of California, Berkeley
Xin Eric Wang, University of California, Santa Barbara and Simular

Reviewers

Rishabh Agrawal, Amazon
Yifei Bi, Facebook
Zhong Chen, Southern Illinois University-Carbondale
Sahil Rajesh Dhayalkar, Brain Corporation
Christopher Driggers-Ellis, University of Florida
Yue Fan, Autodesk
Zhiyuan Gao, University of Southern California
Jingnan Gao, NVIDIA and Shanghai Jiao Tong University
Christan Grant, University of Florida
Jinru Han, University of California, Los Angeles
Patrick Huber, Facebook
Ben Jenkins, Florida Atlantic University
Yanru Jiang, University of California, Los Angeles
Dayeon Ki, Microsoft and University of Maryland, College Park
Sriram Kollipara, Walmart
Pawan Kumar, International Institute of Information Technology Hyderabad
Yuecheng Li, Kuaishou
Chuhan Li, University of California, Santa Barbara
Irene Li, University of Tokyo
Mao Lin, University of California, Merced
Shih-chih Lin, National Tsinghua University
Junbin Lu, University of Washington
Sepideh Mamooler, EPFL
Srijith Ravikumar, Amazon
Anisha Saha, Saarland Informatics Campus, Max-Planck Institute
Tongyue Shi, Peking University
Andrew Shin, Keio University
Kaleen Shrestha, University of Southern California
Thoudam Doren Singh, National Institute of Technology Meghalaya
Sakthivel Sivaraman, NVIDIA
Elior Sulem, Ben-Gurion University of the Negev
Zhimin Sun, AMS AI Lab, Tencent
Rohith Uppala, LinkedIn
Pavan Kumar Velaga, Amazon and Rutgers University
Zirui Wei, C3.ai
Yibo Yan, HKUST and Alibaba
Yang Yan, Southern Illinois University-Carbondale
Junhuan Yang, Amazon

Rongtian Ye, Aalto University
Ziyao Zeng, Yale University
Jing Zhang, Amazon
Pingyue Zhang, Northwestern University
Hang Zhao, Northeastern University
Duo Zhou, University of Illinois at Urbana-Champaign

Table of Contents