MULTIMODAL KDD 2023

INTERNATIONAL WORKSHOP ON MULTIMODAL LEARNING

2023 THEME: MULTIMODAL LEARNING WITH FOUNDATION MODELS
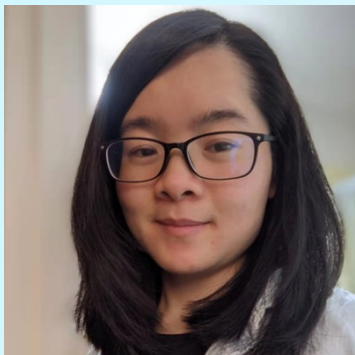
Yuan Ling, Fanyou Wu, Shujing Dong

Yarong Feng, George Karypis, Chandan K Reddy

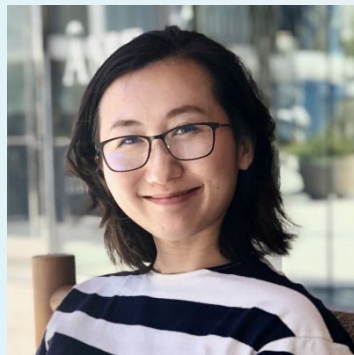Amazon Teams

Aug07,2023, Long Beach, CA

https://multimodal-kdd-2023.github.io/

Multimodal KDD 2023 Workshop

Organizers

Yuan Ling

Fanyou Wu

Shujing Dong

Yarong Feng

George Karypis

Chandan K Reddy

Amazon Teams

Introduction

Multimodal Data	Use Case
Image, text, speech	Dementia Detection
molecular graphs, text-based SMILES encodings	Modeling Olfactory Perception
OCR text, image	Grocery product recognition, stopwatch detection in video
Image, text	E-Commerce Brand Advertising, Open-vocabulary detection, visual question answering, image captioning
Text, Knowledge Graph	Conversational Understanding

Keynotes

Xin (Luna) Dong
Head Scientist
Meta

Keynote Talk 1:
Next-Generation Intelligent Assistants
for AR/VR Devices

Aidong Zhang
Thomas M. Linville Professor of
Computer Science
University of Virginia

Keynote Talk 2:
Graph-based Fusion for
Multimodal Learning

Accepted Papers

Contrastive Multimodal Text Generation for E-Commerce Brand Advertising

Nikhil Madaan, Krishna Kesari, Manisha Verma, Shaunak Mishra and Tor Steiner

Optimizing Learning Across Multimodal Transfer Features for Modeling Olfactory Perception

Daniel Shin, Gao Pei, Priyadarshini Kumari and Tarek Besold

What Makes Good Open-Vocabulary Detector: A Disassembling Perspective

Jincheng Li, Chunyu Xie, Xiaoyu Wu, Bin Wang and Dawei Leng

Explainable Local and Global Models for Fine-Grained Multimodal Product Recognition

Tobias Pettersson, Maria Riveiro and Tuwe Löfström

Evaluating Picture Description Speech for Dementia Detection using Image-text Alignment

Youxiang Zhu, Nana Lin, Xiaohui Liang, John Batsis, Robert Roth and Brian MacWhinney

Graph Meets LLM: A Novel Approach to Collaborative Filtering for Robust Conversational Understanding

Zheng Chen, Ziyan Jiang, Fan Yang, Eunah Cho, Xing Fan, Xiaojiang Huang, Yanbin Lu and Aram Galstyan

Detecting text-rich objects: OCR or object detection? A case study with stopwatch detection

Yarong Feng, Zongyi Liu, Yuan Ling, Shunyan Luo, Shujing Dong, Shuyi Wang and Bruce Ferry

REVEAL: Retrieval-Augmented Visual-Language Pre-Training with Multi-Source Multimodal Knowledge Memory

Ziniu Hu

Schedule

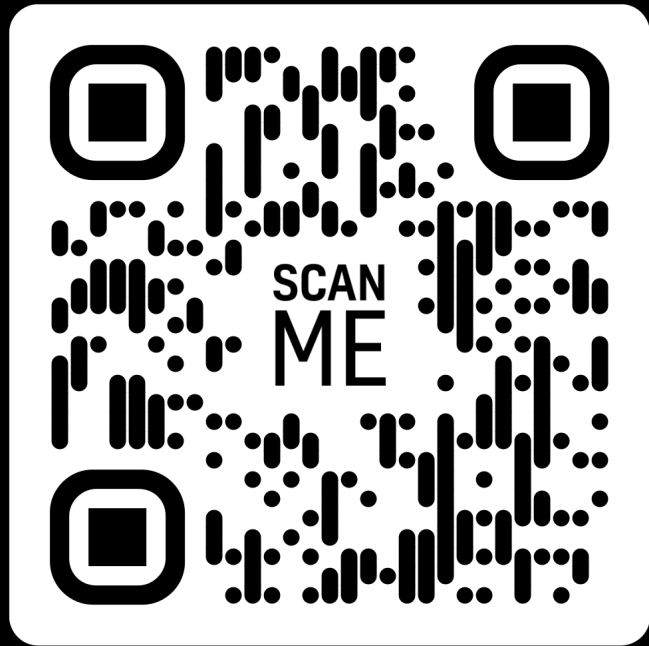| 📖 **Opening** Introduction by organizers.	🕐 1:00 - 1:10 PM
📎 **Keynote Talk 1: Next-Generation Intelligent Assistants for AR/VR Devices** Xin (Luna) Dong *Principal Scientist, Meta*	🕐 1:10 - 1:50 PM
▤ **Evaluating Picture Description Speech for Dementia Detection using Image-text Alignment** Youxiang Zhu, Nana Lin, Xiaohui Liang, John Batsis, Robert Roth and Brian MacWhinney	🕐 1:50 - 2:05 PM
▤ **Optimizing Learning Across Multimodal Transfer Features for Modeling Olfactory Perception** Daniel Shin, Gao Pei, Priyadarshini Kumari and Tarek Besold	🕐 2:05 - 2:20 PM
▤ **Explainable Local and Global Models for Fine-Grained Multimodal Product Recognition** Tobias Pettersson, Maria Riveiro and Tuwe Löfström	🕐 2:20 - 2:35 PM
▤ **Contrastive Multimodal Text Generation for E-Commerce Brand Advertising** Nikhil Madaan, Krishna Kesari, Manisha Verma, Shaunak Mishra and Tor Steiner	🕐 2:35 - 2:50 PM
▤ **What Makes Good Open-Vocabulary Detector: A Disassembling Perspective** Jincheng Li, Chunyu Xie, Xiaoyu Wu, Bin Wang and Dawei Leng	🕐 2:50 - 3:05PM
☕ **Coffee Break**	🕐 3:05 - 3:30 PM
📎 **Keynote Talk 2: Graph-based Fusion for Multimodal Learning** Aidong Zhang *Thomas M. Linville Endowed Professor of Computer Science, University of Virginia*	🕐 3:30 - 4:10 PM
▤ **Detecting text-rich objects: OCR or object detection? A case study with stopwatch detection** Yarong Feng, Zongyi Liu, Yuan Ling, Shunyan Luo, Shujing Dong, Shuyi Wang and Bruce Ferry	🕐 4:10 - 4:25 PM
▤ **REVEAL: Retrieval-Augmented Visual-Language Pre-Training with Multi-Source Multimodal Knowledge Memory** Ziniu Hu	🕐 4:25 - 4:40 PM
▤ **Graph Meets LLM: A Novel Approach to Collaborative Filtering for Robust Conversational Understanding** Zheng Chen, Ziyan Jiang, Fan Yang, Eunah Cho, Xing Fan, Xiaojiang Huang, Yanbin Lu and Aram Galstyan	🕐 4:40 - 4:55 PM
📘 **Closing** Concluding remarks by organizers.	🕐 4:55-5:00 PM

Thank You!

SCAN ME

Multimodal KDD 2023 Workshop

08/07/23

amazon