# Public Seminar Presentation

**By**

**Dr.-Ing. Kumie Gedamu Alemu**

**"Fine-Grained and Multimodal Representation Learning – Challenges and Future Research Directions"**

**Dr.-Ing. Kumie Gedamu Alemu**

## Date and Venue

**Date: 28-Feburary-2025**
**Time: 3:00 pm – 5:00 pm**
**Venue: Virtual**
**Google Meet Link:**
meet.google.com/umi-mizm-axi

I am a research scientist at the Sichuan Artificial Intelligence Research Institute (Yibin), UESTC, China. I was a post-doctoral research fellow in the same group. I obtained my PhD in Computer Science and Technology (Computer Vision) at UESTC, supervised by Prof. Yanli Ji and Prof. Yang Yang. My current research focuses on video understanding, multimodal representation learning approaches.

I have published more than 15 papers in leading AI journals, including IEEE Transactions on Image Processing and Pattern Recognition. I won the Foreign Youth Talent Program and received the honorary title of Sichuan Province's Foreign High-Level Talent in 2024.

My research was funded by the Sichuan science and technology Department, the Ministry of Science and Technology of China, and the National Natural Science Foundation of China's (NSFC) Foreign Young Scholars Research Fund.

## Igniting presentation Abstract

Video understanding stands to benefit significantly from recent advances in fine-grained and multimodal representation learning. This seminar provides an overview of the current state of the art, highlighting both the successes and the remaining challenges. We explore the use of fine-grained representations, multimodal inputs, and joint learning to capture subtle aspects of video features and enhance discriminative feature representation learning. The presentation concludes with a discussion of promising future research directions, including the development of more robust, interpretable, and generalizable models for video understanding and narrative generation.