Multi- and Cross-Modal Information for Enhanced Pattern Recognition

August 21, 2026 - Lyon, France

Co-located with ICPR 2026 (28th International Conference on Pattern Recognition)

SUPPORTED BY

Interreg Italia - Malta

Cofinanziato dall'Unione Europea
Co-funded by the European Union

COESIONE ITALIA 21-27 SICILIA

FONDI.eu

VERA

Welcome to MCMI 2026

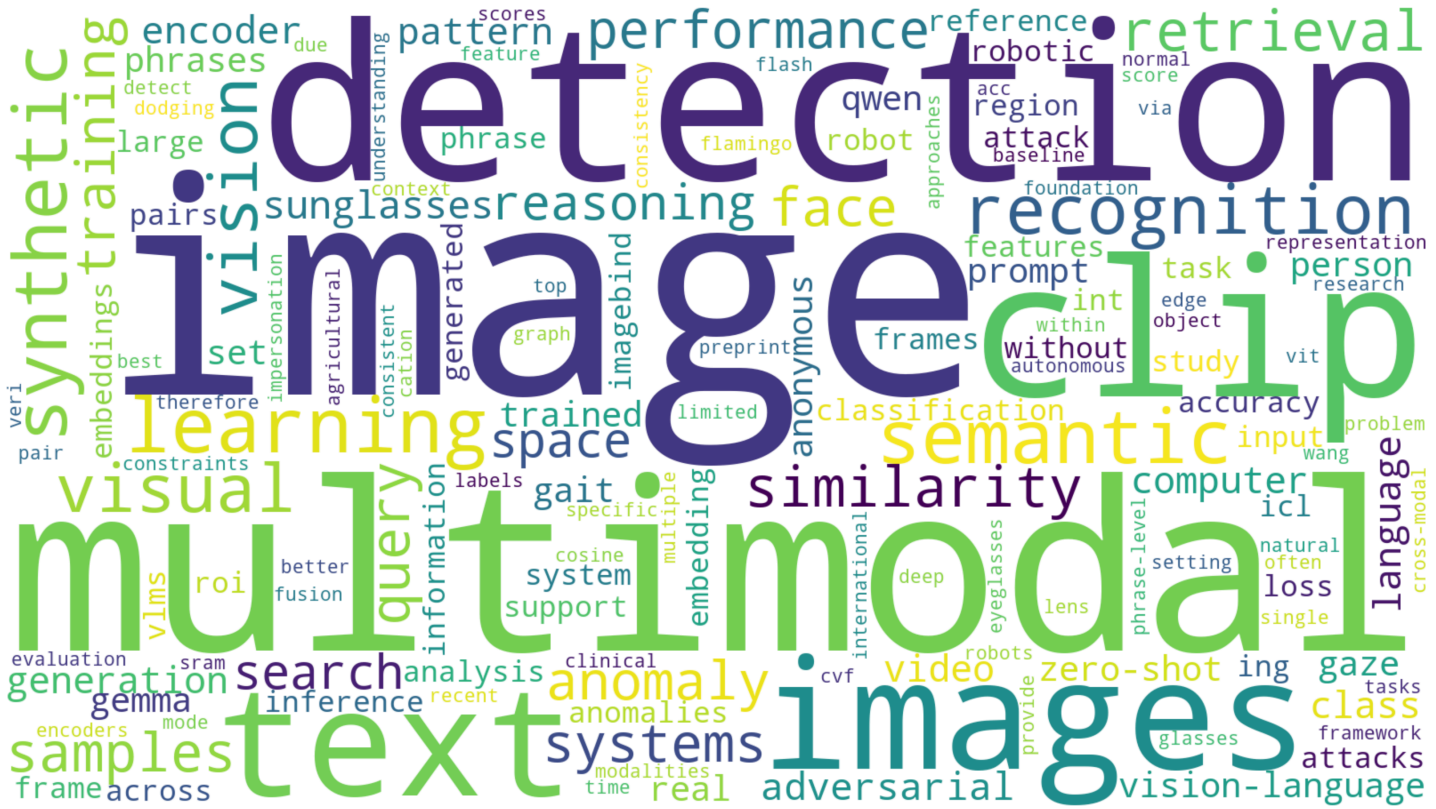
MCMI brings together researchers working across **text, images, audio, and other signals** to improve pattern recognition through multimodal and cross-modal interaction.

- **Fusion methodologies** for combining different modalities
- **Human-machine interaction** grounded in multimodal understanding
- Analysis of **complex, real-world multimodal data**
- Applications in **healthcare**: diagnosis and clinical decision-making

Papers overview

7	3	1
Accepted Papers	Thematic Tracks	Invited Talk

What we will see today



Word cloud of the full text of MCMI 2026 papers

Today's Program

TIME	ITEM
08:30 - 08:40	Welcome & opening remarks
08:40 - 09:20	Keynote: Movement as a Biomarker
09:20 - 10:00	Session 1: Security & Anomaly Detection (2 papers)
10:00 - 10:30	Coffee break
10:30 - 11:30	Session 2: Vision-Language Retrieval (3 papers)
11:30 - 12:10	Session 3: Real-World & Edge Deployment (2 papers)
12:10 - 12:30	Closing remarks & open discussion

Each paper: 15' presentation + 5' Q&A

Keynote

08:40 - 09:20

Movement as a Biomarker: Multimodal Analysis of Handwriting and Gait for Cognitive Disorders

Prof. Francesco Fontanella

Associate Professor, University of Cassino and Southern Lazio, Italy

Session 1: Security & Anomaly Detection

Sunglasses-Region Adversarial Perturbations for Feature-Space Face Recognition Attacks

Chia-Po Wei, Sheng-Pin Chang, Chun-Yao Chen, I-Syuan Lee

Detect, Anticipate, and Explain Anomalies in Agricultural Robots: A Review

Thomas Trouve, Riccardo Bertoglio, Sandro Bimonte, Jocelyn de Goer, Damien Appert

Session 2: Vision-Language Retrieval

SemRank: Semantic Image Retrieval via Knowledge Graph Filtering and CLIP Embeddings

Sourav Pandey, Ashish Singh Patel

Leveraging Phrase-Level Consistency between Images and Texts for Text-Based Person Search

Kota Watanabe, Yuta Shirakawa

Evaluating Pre-trained Image Encoders of Vision-Language and Multimodal Models for Synthetic Facial Image Detection

Liyue Fan, Joseph Roberson

Session 3: Real-World & Edge Deployment

Vision-Language Models for Gait-Based Screening via Zero-Shot and Multimodal In-Context Learning

Moreno La Quatra, Vito Cammarata, Gabriele Trovato, Vincenzo Conti, Valerio Salerno, Salvatore Sorce, Nicole Cilia

Gaze-based Attention-driven Multimodal Single-Object Inference for Microcontroller-Class Edge Vision

Sofia Marilina Glorioso, Andrea Pietro Arena, Ivana Guarneri, Salvatore Sorce

Morning organization

Sessions

Morning only: 08:30 - 12:30

Break

Coffee break: 10:00 - 10:30

Website

mcmi-workshop.github.io/2026

Contact

Email: mcmi@unikore.it

SUPPORTED BY

Interreg Italia - Malta

Cofinanziato dall'Unione Europea
Co-funded by the European Union

COESIONE ITALIA 21-27 SICILIA

FONDI.eu

VERA