

Order	Start Time	End Time	Presentation Title
1	14:00	14:10	Introduction to Workshop-Opening Remarks
2	14:10	14:40	Keynote Speech 1 by Eric Granger
3	14:40	15:30	Oral Session 1
4	15:30	16:00	Keynote Speech 2 by Mohammad Soleymani
5	16:00	16:35	Oral Session 2
6	16:35	16:45	Break
7	16:45	17:15	Keynote Speech 3 by Shaun Canavan
8	17:15	18:00	Poster Session
9	18:00	18:10	Closing Remarks

Oral Session 1

- Introducing Gating and Context into Temporal Action Detection
- MVP: Multimodal Emotion Recognition based on Video and Physiological Signals
- Facial Expression-Enhanced TTS: Combining Face Representation and Emotion Intensity for Adaptive Speech
- TalkinNeRF: Animatable Neural Fields for Full-Body Talking Humans
- Improving face generation quality and prompt following with synthetic captions
- Monitoring Viewer Attention During Online Ads

Oral Session 2

- 7th ABAW Competition: Multi-Task Learning & Compound Expression Recognition
- Textualized and Feature-based Models for Compound Multimodal Emotion Recognition in the Wild
- Compound Expression Recognition via Curriculum Learning
- Affective Behaviour Analysis via Progressive Learning

Poster Session

- Ig3D: Integrating 3D Face Representations in Facial Expression Inference
- Affective Behavior Analysis using Task-adaptive and AU-assisted Graph
- Are We Friends? End-to-End Prediction of Child Rapport in Guided Play
- Better Spanish Emotion Recognition In-the-wild: Bringing Attention to Deep Spectrum Voice Analysis
- Enhancing Facial Expression Recognition through Dual-Direction Attention Mixed Feature Networks: Application to 7th ABAW Challenge
- Boundary Matching and Refinement Network with Cross-modal Contrastive Learning for Temporal Moment Localization
- ABAW7 Challenge: A Facial Affect Recognition Approach based on Transformer Encoder and Multilayer Perceptron
- Smoothing Predictions of Multi-Task EmotiNet Models for Compound Facial Expression Recognition
- Multi-Task Affective Behaviour Analysis based on MT-EmotiNet Models
- Predicting Emotions in Interpersonal Interaction Videos: I Know What You Feel
- MST-KD: Multiple Specialized Teachers Knowledge Distillation for Fair Face Recognition
- rPPG-SysDiaGAN: Systolic-Diastolic Feature Localization in rPPG Using Generative Adversarial Network with Multi-Domain Discriminator
- Tracking Virtual Meetings in the Wild: Re-identification in Multi-Participant Virtual Meetings
- Massively Multi-Person 3D Human Motion Forecasting with Scene Context
- Single Image 3D Human Pose Estimation Using Sequential Joint Group Generation
- Gr-IoU: Ground-Intersection over Union for Robust Multi-Object Tracking with 3D Geometric Constraints
- Are Visual-Language Models Effective Action Recognizer? A Comparative Study