Call for Papers

Special Issue on:

Artificial Intelligence, Affective Computing and Video Analytics for Intelligent Information Systems

The past decade has seen explosive growth in video data (live and on-demand), driven by ubiquitous cameras, mobile devices, and online platforms. From city CCTV and body-cams to retail, sports, tele-health, and social live commerce, organizations now operate end-to-end video streaming pipelines at massive scale (Mao et al., 2024). In parallel, advances in deep learning, spanning video transformers, self-supervised pretraining, and multimodal large models, have transformed what intelligent systems can infer from continuous visual data, moving beyond static clips to long-horizon, low-latency understanding of events, activities, and interactions (Gong et al., 2025).

Affective computing adds a new dimension by interpreting emotional and social signals in video and multimedia. For example, research shows that portraying positive emotions and trust cues in video significantly increases viewer engagement. By combining facial expression, gesture, and acoustic analysis, advanced AI can infer user sentiment and intent. These capabilities enrich social-media intelligence, recommender systems, and customer journey analysis.

This special issue seeks cutting-edge work at the intersection of video analytics and affective computing, harnessing emotion and intent cues in visual data to enhance information systems, and about AI for video streaming: methods, systems, and governance for real-time ingestion, analysis, and decisioning over continuous video flows.

We aim to bridge computer-vision analytics and affective models (extracting sentiment, engagement, trustworthiness, etc.) to design intelligent, human-centered systems that inform decision-making and adaptive services We are interested in methods that fuse visual, auditory, and textual data (multimodal analysis) to detect emotion, sentiment, engagement, gaze, gesture, posture, prosody and language in real time. Such techniques, for example, enable insights from live-streamed user interactions or “experience mining” in marketing and e-commerce (Stappen et al., 2022). Deep-learning advances enable machines to recognize facial expressions, gestures, tone of voice and other cues to infer sentiment, engagement and intent. Recent work demonstrates, for instance, that Convolutional Neural Networks combining video and audio input can predict viewers’ emotions with high accuracy, achieving around a 75% AUC on real-world video-ads data (Antonov et al., 2024). Other research has integrated language and vision: one study built an “Affective Mimicry Index” from CEO interview videos, using computer vision together with large language models , and found that higher video-derived empathy scores correlate with better firm performance (Cui et al., 2025). These examples show how multimodal video analytics can extract rich affective signals, from basic expressions to higher-level trust and empathy cues. Our goal is to bring these advances into intelligent information systems that are explainable, fair, and aligned with human values.

We invite submissions of original research, surveys and perspectives that explore the intersection of video analytics, affective computing and intelligent systems design. Topics of interest include (but are not limited to):

- Scalable video understanding: algorithms and architectures for analyzing large-scale video streams (e.g. in surveillance, social media or enterprise) to support decision-making and business analytics.
- Emotion and sentiment analysis: detection of affective and engagement cues in video (e.g. customer-journey recordings, advertisement viewing, user-generated content).
- Multimodal fusion: methods that combine vision with audio, text or other modalities to enrich social-media intelligence and contextual analysis.
- Streaming-native AI and MLOps: online inference under tight latency budgets, stream processing/windowing, drift detection, A/B testing and safe rollouts, cost–latency–accuracy trade-offs, and end-to-end observability for production pipelines.
- Video(-language) models for streaming: multimodal LLMs/VLMs, memory and token compression for long-horizon video, retrieval-augmented streaming
- Real-time and distributed architectures: edge/cloud or federated systems for real-time video analytics at scale, including considerations of latency, bandwidth and privacy.
- Explainability, fairness and compliance: approaches to make affective video AI transparent and trustworthy, addressing ethical, legal and regulatory challenges.
- Adaptive learning: techniques for domain adaptation, continual learning or few-shot learning to handle evolving video streams and reduce the need for large labeled datasets.
- Case studies and applications: empirical studies in domains such as healthcare (e.g. remote patient monitoring), finance (e.g. behavioral analytics), education (e.g. XR/serious games environments), and smart environments.

Forms of Submission:

This special issue will consist of:
(1) the best submissions from an open call for papers selected on a competitive basis
(2) invited submissions that are extended or modified versions of selected papers accepted at the Workshop 'Extended Reality and Serious Games for Education and Learning' (ERSeGEL, https://csedu.scitevents.org/ERSeGEL.aspx), part of the **18th International Conference on Computer Supported Education** (CSEDU 2026, https://csedu.scitevents.org/, May 18-20, 2026) and at the International Conference on eXtended Reality (XR Salento 2026, https://xrsalento.it/). **In both cases, only papers focused on research topics related to this special issue will be considered**. The submissions will be substantial revisions of the two conferences publications, and the authors will be required to submit a letter detailing the difference between the conference paper and the new version.

Submitted manuscripts should not have been published previously, nor be under consideration for publication elsewhere (except conference proceedings papers). Conference papers should be cited and noted on the paper. Please note that submitted extended papers should contain at least 40% new content (e.g., in the form of technical extensions, more in-depth evaluations, or additional use cases) and not exceed 30% copy/paste from the conference paper.

Every submission will be evaluated by at least two independent reviewers, using the review process and standards of the Information Systems Frontiers journal for regular submissions. The Editor-in-Chief, along with the guest editors, will make the final decision to accept or reject a submission, based on the reviews.

Submission Instruction

Manuscripts must be submitted in PDF format to the ISF-Springer online submission system at https://www.editorialmanager.com/isfi and the authors need to select "Special Issue: Video Analytics and Affective Computing for Intelligent Information Systems" during the submission process. Paper submissions must conform to the format guidelines of Information Systems Frontiers available at https://www.springer.com/journal/10796/submission-guidelines.

Guest Editors:

- Dr. Andrea Generosi, Department of Information Science and Technology, Pegaso University | Email: andrea.generosi@unipegaso.it
- Prof. Luigi Gallo: Department of Information Science and Technology, Pegaso University | Email: luigi.gallo@unipegaso.it
- Dr. Valerio De Luca, Department of Information Science and Technology, Pegaso University | Email: valerio.deluca@unipegaso.it
- Prof. Maura Mengoni, Department of Industrial Engineering and Mathematical Sciences, Marche Polytechnic University | Email: m.mengoni@univpm.it
- Dr. Josef Spjut, NVIDIA | Email: jspjut@nvidia.com

Important dates:

- Submission of abstract (up to 300 words): May 30, 2026
- Notification of abstract acceptance: June 15, 2026
- Submission of full paper: July 20, 2026
- Notification of first round reviews: October 15, 2026
- Revised manuscript due: November 15, 2026
- Notification of second round reviews: December 20, 2026
- Final version due: January 31, 2027
- Tentative publication date: Late February 2027

Guest Editors: Biographies

Dr. Andrea Generosi: Tenure-Track Assistant Professor with experience in the design and development of innovative solutions based on Deep Learning. PhD in Industrial Engineering, engaged in research and development projects for industry and the cultural sector, with a focus on Affective Computing, ergonomics, automotive, HCI and Extended Reality. Previously a research fellow and adjunct professor at the Polytechnic University of Marche and co-founder of the UNIVPM Emoj Srl startup and spin-off.

Prof. Luigi Gallo: Associate Researcher at CNR-ICAR since 2024, Full Professor of Information Processing Systems at Pegaso University and Director of the xRAI Research Centre for Computer Vision and XR. He coordinates national and international ICT projects for health and cultural heritage. Associate Editor of Pattern Recognition Letters and Heritage, and organiser of IEEE SITIS, AICCSA, and XR Salento. His research spans HCI, computer vision, pattern recognition, and XR, with 130+ publications.

Dr. Valerio De Luca: Tenure-Track Assistant Professor specializing in extended reality for medicine, education, and cultural heritage. Past work includes grid computing, real-time AV streaming, and QoE evaluation. Experienced in GPU computing for UAV path planning and AR to enhance situational awareness in remote drone control. His recent research focuses on extended reality for preoperative planning and intraoperative support in surgery and on human-computer interaction in extended realtity applications

Prof. Maura Mengoni: Associate Professor of Design and Methods for Industrial Engineering at the Marche Polytechnic University, where she leads the Virtual Prototyping Laboratory. Her scientific activity, documented by over 180 publications and more than 90 national and international research projects, focuses on industrial design, human–machine interaction, extended reality (XR) technologies, and artificial intelligence for human and emotional factor analysis. She is co-founder and scientific advisor of the university spin-off EMOJ S.r.l., an innovative startup specializing in AI-driven emotional analytics, and collaborates with several European research centers and universities on inclusive design, ergonomics, and immersive technologies.

Dr. Josef Spjut is a Senior Research Scientist at NVIDIA working on human-computer experiences, such as esports and augmented reality. His research contributed to the ray tracing hardware in the Turing architecture (RTX 20 series) and the NVIDIA Reflex esports platform launched alongside Ampere (RTX 30 series). Previously he was a Visiting Assistant Professor in the department of Engineering at Harvey Mudd College. He received a Ph.D. from the University of Utah and a B.S. from University of California, Riverside, both in Computer Engineering

References:

Mao, M., Lee, A., & Hong, M. (2024). Deep learning innovations in video classification: a survey on techniques and dataset evaluations. *Electronics*, *13*(14), 2732. https://doi.org/10.3390/electronics13142732

Gong, L., Yang, H., Fang, G., Ju, B., Guo, J., Zhu, X., ... & Boukerche, A. (2025). A Survey on Video Analytics in Cloud-Edge-Terminal Collaborative Systems. *arXiv preprint arXiv:2502.06581*.

Stappen, L., Baird, A., Lienhart, M., Bätz, A., & Schuller, B. (2022). An estimation of online video user engagement from features of time-and value-continuous, dimensional emotions. *Frontiers in Computer Science*, *4*, 773154. https://doi.org/10.3389/fcomp.2022.773154

Antonov, A., Kumar, S. S., Wei, J., Headley, W., Wood, O., & Montana, G. (2024). Decoding viewer emotions in video ads. Scientific Reports, 14(1), 26382. https://doi.org/10.1038/s41598-024-76968-9

Cui, Li and Lu, Zhoudao and Ng, Ka Chung and Zhao, Jingran, Can Machines Read Your Face? A Video Analytics Framework for Measuring Empathy via Emotional Mimicry in Video Recordings (May 19, 2025). http://dx.doi.org/10.2139/ssrn.5260163