

1. Record Nr.	UNISA996641267303316
Autore	Ide Ichiro
Titolo	MultiMedia Modeling : 31st International Conference on Multimedia Modeling, MMM 2025, Nara, Japan, January 8–10, 2025, Proceedings, Part IV // edited by Ichiro Ide, Ioannis Kompatsiaris, Changsheng Xu, Keiji Yanai, Wei-Ta Chu, Naoko Nitta, Michael Riegler, Toshihiko Yamasaki
Pubbl/distr/stampa	Singapore : , : Springer Nature Singapore : , : Imprint : Springer, , 2025
ISBN	9789819620715 9819620716
Edizione	[1st ed. 2025.]
Descrizione fisica	1 online resource (485 pages)
Collana	Lecture Notes in Computer Science, , 1611-3349 ; ; 15523
Altri autori (Persone)	KompatsiarisYiannis XuChangsheng YanaiKeiji ChuWei-Ta NittaNaoko RieglerMichael YamasakiToshihiko
Disciplina	006.37
Soggetti	Computer vision Image processing - Digital techniques Signal processing Pattern recognition systems Application software Information storage and retrieval systems Computer Vision Computer Imaging, Vision, Pattern Recognition and Graphics Signal, Speech and Image Processing Automated Pattern Recognition Computer and Information Systems Applications Information Storage and Retrieval
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia

Regular Papers -- SES-Net: Multi-dimensional Spot-Edge-Surface Network for Nuclei Segmentation -- Skin-Adapter: Fine-Grained Skin-Color Preservation for Text-to-Image Generation -- Small Tunes Transformer: Exploring Macro & Micro-Level Hierarchies for Skeleton-Conditioned Melody Generation -- SMG-Diff: Adversarial Attack Method Based on Semantic Mask-Guided Diffusion -- SPLGAN-TTS: Learning Semantic and Prosody to Enhance the Text-to-Speech Quality of Lightweight GAN Models -- SSCDUF: Spatial-Spectral Correlation Transformer Based on Deep Unfolding Framework for Hyperspectral Image Reconstruction -- SSDL: Sensor-to-Skeleton Diffusion Model with Lipschitz Regularization for Human Activity Recognition -- Structural Information-guided Fine-grained Texture Image Inpainting -- Style Separation and Content Recovery for Generalizable Sketch Re-identification and A New Benchmark -- Synchronization and Calibration of Video Sequences acquired using Multiple Plenoptic 2.0 Cameras -- Target-Oriented Dynamic Denosing Curriculum Learning for Multimodal Stance Detection -- TDM: Temporally-Consistent Diffusion Model for All-in-One Real-World Video Restoration -- Temporal Closeness for Enhanced Cross-Modal Retrieval of Sensor and Image Data -- The Right to an Explanation under the GDPR and the AI Act -- Toward Appearance-based, Autonomous Landing Site Identification for Multirotor Drones in Unstructured Environments -- Towards Inclusive Education: Multimodal Classification of Textbook Images for Accessibility -- Towards Visual Storytelling by Understanding Narrative Context through Scene-Graphs -- TPS-YOLO: The Efficient Tiny Person Detection Network Based on Improved YOLOv8 and Model Pruning -- Uncertainty-guided Joint Semi-supervised Segmentation and Registration of Cardiac Images -- Understanding the Roles of Visual Modality in Multimodal Dialogue: An Empirical Study -- Vision-Language Pretraining for Variable-shot Image Classification -- Visual Anomaly Detection on Topological Connectivity under Improved YOLOv8 -- Wavelet Integrated Convolutional Neural Network for ECG Signal Denoising -- WavFusion: Towards wav2vec 2.0 Multimodal Speech Emotion Recognition -- Zero-shot Sketch-based Image Retrieval with Hybrid Information Fusion and Sample Relationship Modeling -- Special Session: ExpertSUM: Special Session on Expert-Level Text Summarization from Fine-Grained Multimedia Analytics -- CalorieVoL: Integrating Volumetric Context into Multimodal Large Language Models for Image-based Calorie Estimation -- Can Masking Background and Object Reduce Static Bias for Zero-shot Action Recognition? -- Special Session: MLLMA: Special Session on Multimodal Large Language Models and Applications -- Enhanced Anomaly Detection in 3D Motion through Language-Inspired Occlusion-Aware Modeling -- Evaluating VQA Models' Consistency in the Scientific Domain -- Image2Text2Image: A Novel Framework for Label-Free Evaluation of Image-to-Text Generation with Text-to-Image Diffusion Models -- Quantifying Image-Adjective Associations by Leveraging Large-Scale Pretrained Models -- TACST: Time-Aware Transformer for Robust Speech Emotion Recognition -- TS-MEFM: A New Multimodal Speech Emotion Recognition Network Based on Speech and Text Fusion.

This five-volume set LNCS 15520-15524 constitutes the proceedings of the 31st International Conference on Multimedia Modeling, MMM 2025, held in Nara, Japan, January 8–10, 2025. The 135 full papers and 41 short papers presented in these proceedings were carefully reviewed and selected from 348 submissions. The MMM conference was organized in topics related to multimedia modelling, particularly: audio, image, video processing, coding and compression; multimodal

analysis for retrieval applications, and multimedia fusion methods.
