

1. Record Nr.	UNINA9910806191303321
Autore	Rudinac Stevan
Titolo	MultiMedia Modeling [[electronic resource]] : 30th International Conference, MMM 2024, Amsterdam, The Netherlands, January 29 – February 2, 2024, Proceedings, Part II // edited by Stevan Rudinac, Alan Hanjalic, Cynthia Liem, Marcel Worring, Björn Þór Jónsson, Bei Liu, Yoko Yamakata
Pubbl/distr/stampa	Cham : , : Springer Nature Switzerland : , : Imprint : Springer, , 2024
ISBN	3-031-53308-9
Edizione	[1st ed. 2024.]
Descrizione fisica	1 online resource (540 pages)
Collana	Lecture Notes in Computer Science, , 1611-3349 ; ; 14555
Altri autori (Persone)	HanjalicAlan LiemCynthia WorringMarcel JónssonBjö Þór LiuBei YamakataYoko
Disciplina	006.37
Soggetti	Computer vision Image processing Pattern recognition systems Application software Information storage and retrieval systems Machine learning Computer Vision Image Processing Automated Pattern Recognition Computer and Information Systems Applications Information Storage and Retrieval Machine Learning
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Nota di contenuto	Self-distillation Enhanced Vertical Wavelet Spatial Attention for Person Re-identification -- High Capacity Reversible Data Hiding in Encrypted

Images Based on Pixel Value Preprocessing and Block Classification -- HPAttack: An Effective Adversarial Attack for Human Parsing -- Dynamic-Static Graph Convolutional Network for Video-Based Facial Expression Recognition -- Hierarchical Supervised Contrastive Learning for Multimodal Sentiment Analysis -- Semantic Importance-Based Deep Image Compression Using A Generative Approach -- Drive-CLIP: Cross-modal Contrastive Safety-Critical Driving Scenario Representation Learning and Zero-shot Driving Risk Analysis -- MRHF: Multi-stage Retrieval and Hierarchical Fusion for Textbook Question Answering -- Multi-scale Decomposition Dehazing with Polarimetric Vision -- CLF-Net: A Few-shot Cross-Language Font Generation Method -- Multi-dimensional Fusion and Consistency for Semi-supervised Medical Image Segmentation -- Audio-Visual Segmentation By Leveraging Multi-Scaled Features Learning -- Multi-head Hashing with Orthogonal Decomposition for Cross-modal Retrieval -- Fusion Boundary and Gradient Enhancement Network for Camouflage Object Detection -- Find the Cliffhanger: Multi-Modal Trailerness in Soap Operas -- SM-GAN: Single-stage and Multi-object Text Guided Image Editing -- MAVAR-SE: Multi-scale Audio-Visual Association Representation Network for End-to-end Speaker Extraction -- NearbyPatchCL: Leveraging Nearby Patches for Self-Supervised Patch-Level Multi-Class Classification in Whole-Slide Images -- Improving Small License Plate Detection with Bidirectional Vehicle-plate Relation -- A Purified Stacking Ensemble Framework for Cytology Classification -- SEAS-Net: Segment Exchange Augmentation for Semi-Supervised Brain Tumor Segmentation -- Super-Resolution-Assisted Feature Refined Extraction for Small Objects in Remote Sensing Images -- Lightweight Image Captioning Model Based on Knowledge Distillation -- Irregular License Plate Recognition via Global Information Integration -- TNT-Net: Point Cloud Completion by Transformer in Transformer -- Fourier Transformer for Joint Super-Resolution and Reconstruction of Mr Image -- MVD-NeRF: Resolving Shape-Radiance Ambiguity via Mitigating View Dependency -- DPM-Det: Diffusion Model Object Detection Based on DPM-Solver++ Guided Sampling -- CT-MVSNet: Efficient Multi-View Stereo with Cross-scale Transformer -- A Coarse and Fine Grained Masking Approach for Video-grounded Dialogue -- Deep self-supervised subspace clustering with triple loss -- LigCDnet: Remote Sensing Image Cloud Detection Based on Lightweight Framework -- Gait Recognition Based on Temporal Gait Information Enhancing -- Learning Complementary Instance Representation with Parallel Adaptive Graph-Based Network for Action Detection -- CEGSegNet: Context-Enhancement Semantic Segmentation Network Based on Transformer -- MoCap-Video Data Retrieval with Deep Cross-Modal Learning -- LRATNet: Local-Relationship-Aware Transformer Network for Table Structure Recognition.

Sommario/riassunto

This book constitutes the refereed proceedings of the 30th International Conference on MultiMedia Modeling, MMM 2024, held in Amsterdam, The Netherlands, during January 29–February 2, 2024. The 112 full papers included in this volume were carefully reviewed and selected from 297 submissions. The MMM conference were organized in topics related to multimedia modelling, particularly: audio, image, video processing, coding and compression; multimodal analysis for retrieval applications, and multimedia fusion methods.
