ACM International Conference on Multimedia in Asia

Program Guide

Hybrid/Tokyo, Japan
13-16 December, 2022

MESSAGE FROM THE GENERAL AND PROGRAM CHAIRS

On behalf of the Organizing Committee, it is our great pleasure to welcome you to the 4th ACM International Conference on Multimedia in Asia (MM Asia 2022). The 4th ACM Multimedia Asia is held in Tokyo, Japan with an extensive program that includes technical sessions covering all aspects of the multimedia field in forms of regular, short, demo papers, keynotes, tutorials and workshops. Due to the COVID-19 pandemic, it is held in a hybrid mode by offering both onsite and online events. Conference registrants have full access to all the above activities in the program via either onsite or online.

The ACM Multimedia Asia conference series was established in 2019 by putting together the long-lasting experience of former PCM and ICIMCS, which both have good history as well as attending experiences. Officially sponsored by ACM SIGMM, MM Asia is a newly established international conference to showcase the scientific achievements and industrial innovations in the multimedia field. Its mission is to illuminate the state of the art in multimedia computing by bringing together researchers and practitioners in this field.

This year, we received 85 submissions to the regular paper track, short paper track, and demo paper track. These submissions cover a wide range of research topics, including multimedia and vision, multimodal analysis and description, multimedia search and recommendation, music and speech processing in multimedia, social multimedia, multimedia virtual and augmented reality, multimedia arts, entertainment and culture, and multimedia for healthcare. We thank the 109 Technical Program Committee members who spent many efforts reviewing papers and providing valuable and constructive feedback on the papers. Ninety-four percent of the papers were assigned to three reviewers and the remainder to two reviewers, with the Program Chairs carefully reading the papers in the case of the two reviewers. From the papers submitted to the regular paper track, the program chairs decided to accept 27 oral papers (40.3%). Six papers of high quality in the remainder were also selected for inclusion in the short paper track. For the short paper track, 5 submissions were accepted as poster papers, and for the demo track, 9 submissions were accepted as demo papers. Among the 27 oral papers, five papers were selected as the best paper candidate, to compete for the Best Paper and the Best Paper RunnerUp awards. In addition, two tutorials: "Synthetic Data and Multimedia" and "Human-centric Visual Understanding" and workshops: "Multimedia Understanding with Pre-trained Models" are held in MM Asia 2022. With all the efforts and help, we are confident to say that MM Asia 2022 will enable you to enjoy an outstanding program, and exchange your ideas with leading researchers in various disciplines of multimedia. We sincerely thank all the authors for their recognition and support of our conference.

We would like to express our sincere gratitude to all the organizing committee members, to help organize all of the sessions. Their contributions are much appreciated. It is their outstanding effort in preparing this rich and complex program that characterizes the 4th ACM Multimedia Asia conference. We would like to appreciate the financial support from the foundations including Kajima Foundation, Tateisi Science and Technology Foundation, and the Telecommunications Advancement Foundation: Their grants for holding the

international conference are very important for us to reduce the author registration fee and to waive the online student attendee fee.

We sincerely hope that you will enjoy your stay in Tokyo or your virtual experience and value your participation in MM Asia 2022.

GENERAL CO-CHAIRS

Keiji Yanai	The University of Electro-Communications
Kiyoharu Aizawa	The University of Tokyo
Phoebe Chen	La Trobe
Shuqiang Jiang	CAS

PROGRAM CO-CHAIRS

Rei Kawakami	Titec
Toshihiko Yamasaki	The University of Tokyo
Winston Hsu	NTU
Jing Liu	CAS
Wei Tsang Ooi	NUS

ORGANIZERS

GENERAL CHAIRS

Kiyoharu Aizawa	The University of Tokyo
Phoebe Chen	La Trobe
Shuqiang Jiang	CAS
Keiji Yanai	The University of Electro-Communications

PROGRAM CHAIRS

Toshihiko Yamasaki	The University of Tokyo
Jing Liu	CAS
Winston Hsu	NTU
Wei Tsang Ooi	NUS
Rei Kawakami	Titech

LOCAL CHAIRS

Yusuke Matsui	The University of Tokyo
Xueting Wang	CyberAgent
Zheng Wang	Wuhan University

PUBLICITY CHAIRS

Shin'ichi Satoh	NII
Wen-Huang Cheng	National Yang Ming Chao Tung University
Vincent Oria	New Jersey Institute of Technology
Klaus Schöffmann	Klagenfurt University
Changsheng Xu	CAS
Shintami Chusnul Hidayati	Institut Teknologi Sepuluh Nopember
Huynh Thi Thanh Binh	Hanoi University of Science and Technology
Yong Man Ro	KAIST

PANEL CHAIRS

Jianquan Liu	NEC
Mohan Kankanhalli	NUS
Balakrishnan Prabhakaran	University of Texas at Dallas

TUTORIAL CHAIRS

Naoko Nitta	Osaka University
Weiqing Min	CAS

PROGRAM COMMITTEE MEMBERS

Alan Smeaton	Insight Centre for Data Analytics, Dublin City University
Atsushi Hashimoto	OMRON SINIC X Corp.
Barry Theobald	Apple
Bjorn Stenger	Rakuten Institute of Technology
Calvin Ku	National Tsinghua University
Carl Debono	University of Malta
Chuanbin Liu	USTC
Deborah Dahl	Conversational Technologies
Dong Zhang	Soochow University
Fan Liu	National University of Singapore
Gang Wu	Adobe Research
Guan-Ming Su	Dolby Labs
Guo Zhong	Guangdong University of Foreign Studies
Haichao Yu	UIUC
Hanchao Yu	UIUC
Hehe Fan	National University of Singapore
Hikaru Ikuta	The University of Tokyo
Hisashi Miyamori	Kyoto Sangyo University
Huaiwen Zhang	NLPR
Huanglu Wen	Beijing Institute of Technology
Hui Cui	Shandong Normal Unversity
Huiyuan Yang	Rice University
Humera Tariq	University of Karachi
Ichiro Ide	Nagoya University
Jiaming Zhang	Beijing Jiaotong University, China
Jianbo Wang	The University of Tokyo
Jian-Wei Peng	National Cheng Kung University
Jianxin Li	Beihang University
Jie Cao	Institute of Automation, Chinese Academy of Sciences
Jingwen Chen	JD AI Research
Jiwei Wei	University of Electronic Science and Technology of China
John Kender	Columbia University
John See	Heriot-Watt University Malaysia
Junbin Xiao	National University of Singapore
Junwei Liang	Carnegie Mellon University
Jun-Yan He	DAMO Academy, Alibaba Group
Kaixi Hu	Wuhan University of Technology
Kimiaki Shirahama	Kindai University

Kohei Matsuzaki	KDDI Research
Koki Tsubota	The University of Tokyo
Lai-Kuan Wong	Multimedia University
Lang Huang	The University of Tokyo
Ling Xiao	The University of Tokyo
Luca Rossetto	University of Zurich
Luwei Zhang	The University of Tokyo
Marc A. Kastner	Kyoto University
Marc Cavazza	National Institute of Informatics
Na Zheng	National University of Singapore
Nakamasa Inoue	Tokyo Institute of Technology
Nelson Yalta	Hitachi RDG
Oscar Saz	Emotech Ltd
Palaiahnakote Shivakumara	University of Malaya
Peng Zhou	Anhui University
Pengpeng Zhao	Soochow University
Qinqin Zhou	Xiamen University
Ruifan Li	Beijing University of Posts and Telecommunications
Ryota Yoshihashi	Tokyo Institute of Technology
Sarala Padi	NIST
Satoshi Kosugi	The University of Tokyo
Shan Xue	CSIRO
Sheng Li	National Institute of Information & Communications Technology (NICT)
Shengzhou Yi	The University of Tokyo
Shih-Wei Sun	Taipei National University of the Arts
Shuhui Jiang	Northeastern University
Sicheng Zhao	Columbia University
Siyu Xia	Southeast University, China
Songqing Chen	George Mason University
Sourav Mishra	The University of Tokyo
Takashi Shibata	NTT Corporation
Tao Chen	Nanjing University of Science and Technology
Tao He	Monash University
Tao Yao	Ludong University
Trung Hieu Nguyen	Alibaba Cloud Singapore
Tse-Yu Pan	National Tsing Hua University
Wei Zhang	Singapore Institute of Technology
Weifeng Liu	China University of Petroleum (East China)
Weiqing Min	Institute of Computing Technology, Chinese Academy of Sciences
Weixin Li	Beihang University
Wen-Cheng Chen	National Cheng Kung University
Werner Bailer	JOANNEUM RESEARCH
Xiangbo Shu	Nanjing University of Science and Technology

Xiangteng He	Peking University
Xiangyang Li	Institute of Computing technology, CAS
Xiao Wang	Wuhan University
Xiaoheng Jiang	Zhengzhou Universtiy
Xin Luo	Shandong University
Xing Wei	Xi'an Jiaotong University
Xingbo Liu	Shandong Jianzhu University
Xin-Shun Xu	Shandong University
Yanbin Hao	University of Science and Technology of China
Yang Li	East China Normal University
Yangyang Guo	National University of Singapore
Yi Jiang	Bytedance
Yi Zhang	Beijing Jiaotong University
Ying Fu	Beijing Institute of Technology
Yiwei Zhang	The University of Tokyo
Yoko Yamakata	University of Tokyo, Japan
Yu Sun	Harbin Institute of technology
Yuansong Qiao	Athlone Institute of Technology
Yujuan Ding	The Hong Kong Polytechnic University
Yukinobu Taniguchi	Tokyo University of Science
Yunlu Chen	University of Amsterdam
Yuxin Peng	Peking University
Zeren Sun	Nanjing University of Science and Technology
Zhe Liu	Huazhong University of Science and Technology
Zhifan Gao	Sun Yat-sen University
Zhilin Zhang	Amazon
Zhiyong Cheng	Shandong Academy of Sciences
Zichun Zhong	Wayne State University

PROGRAM AT A GLANCE

December

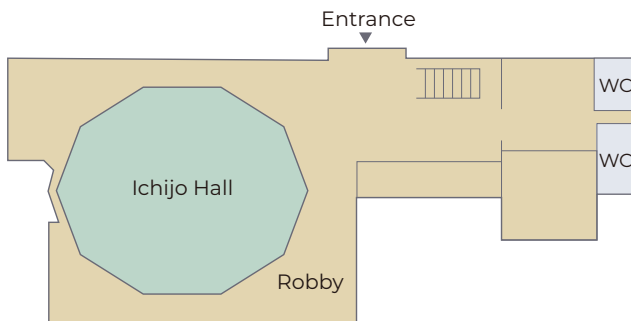| 13	14	15	16
09:15 **Registration** 09:45 – 11:45 (Ichijo Hall) **Workshop** Multimedia Understanding with Pre-trained Models	09:45 – 10:00 (Ichijo Hall) **Opening** 10:00 – 11:40 (Ichijo Hall) **Oral 1-1** Award Session	10:00 – 11:30 (Ichijo Hall) **Oral 2-1** Video Compression, Broadcasting, and Analysis	10:00 – 11:30 (Ichijo Hall) **Oral 3-1** Low-level Vision and Image Processing
13:00 – 15:00 (Ichijo Hall) **Tutorial 1** Synthetic Data and Multimedia 15:00-17:00 (Ichijo Hall) **Tutorial 2** Human-centric Visual Understanding	13:00 – 14:00 (Ichijo Hall) **Keynote 1** Machine Learning for Creative Workflow 14:00 – 14:30 (Ichijo Hall) **Demo Spotlight** 14:30 – 15:45 (Lobby) **Poster+Demo** 16:00 – 17:00 (Ichijo Hall) **Oral 1-2** Text, Speech, and Vision	13:00 – 14:00 (Ichijo Hall) **Keynote 2** Connecting the Dots: Digital Humanities and Historical Big Data Research for Japanese Culture 14:00 – 15:00 (Ichijo Hall) **Short Spotlight** 15:15 – 17:00 (Lobby) **Poster+Demo**	13:00 – 14:30 (Ichijo Hall) **Oral 3-2** Robustness, Data Augmentation and Disentangling 14:30 – 14:45 (Ichijo Hall) **Closing**

Entrance

Ichijo Hall

Robby

WC

WC

KEYNOTE TALKS

Machine Learning for Creative Workflow

December 14, 13:00 - 14:00

Kota Yamaguchi is a research manager at CyberAgent AI Lab. He currently works on research and development of computer vision and machine learning techniques for creative workflow automation. He was previously an assistant professor at Tohoku University from 2014 to 2017. He received a Ph.D. degree in Computer Science from Stony Brook University in 2014. He received an MS in 2008 and a BE in 2006, both from the University of Tokyo.

Kotaro Yamaguchi
CyberAgent, Inc.

Abstract

Recent advancement of machine learning techniques transforms how creators design and finish up their work. In this talk, I present ongoing research efforts at CyberAgent in bringing the machine learning techniques to intelligently assist graphic designers in the digital advertising industry. The data structure around creative workflow is characterized by a vector graphic format that precisely describes the high-level multi-modal structure of the final visual presentation, which poses several challenges to machine learning models from data representation to performance evaluation. We will discuss how we approach design tasks such as text de-rendering from raster images and unsupervised document generation.

Connecting the Dots: Digital Humanities and Historical Big Data Research for Japanese Culture

December 15, 13:00 - 14:00

Asanobu Kitamoto
NII

Asanobu Kitamoto earned his Ph.D. in electronic engineering from the University of Tokyo in 1997. He is now the Director of the Center for Open Data in the Humanities (CODH), Joint Support-Center for Data Science Research (DS), Research Organization of Information and Systems (ROIS), Professor of the National Institute of Informatics, and SOKENDAI (The Graduate University for Advanced Studies). He has developed various data-driven science approaches in fields such as the humanities, earth sciences, and disaster management. He has released databases and software as academic research platforms with a few million users from academia and society. He is also working on a trans-disciplinary collaboration to promote open science. He has received awards such as Jury Recommended Works (Art Division) from Japan Media Arts Festival, Yamashita SIG Research Award from the Information Processing Society of Japan (IPSJ), Best Paper Award from IPSJ SIG Computers and the Humanities Symposium, Academic Award (Research Paper) from Japan Society for Digital Archive, and Good Design Award.

Abstract

Data-driven approaches such as machine learning and multimedia technology can accelerate research on Japanese culture from a historical perspective. Furthermore, thanks to a widespread movement toward open science, such as open data and open source, research on Japanese culture has finally entered into a big data era. However, humanities data is full of interpretation and meaning, has a complex structure with diversity, and requires implicit knowledge not explicitly described inside the humanities data. Hence we do not expect that a technological silver bullet, such as AI, can solve all the problems at once. Instead, we employ a connect-the-dot approach by developing and connecting tools and datasets to answer humanities research questions for understanding Japanese culture. For that purpose, we carefully gather evidence from many sources in the past, create structured data with human-machine collaborations, and integrate them as linked data to draw a bigger picture of Japanese culture looking backward from the present. The talk will introduce our activities in the ROIS-DS Center for Open Data (CODH) in the Humanities to tackle these challenges as digital humanities and historical big data research. The talk will include machine learning for the recognition of Kuzushiji (Japanese historical cursive scripts), art history research using IIIF (International Image Interoperability Framework), and computer vision-based differential reading for diachronic transcription. We also discuss the importance of domain knowledge and collaboration with domain experts to ask meaningful research questions beyond simplistic metric-based evaluations.

TUTORIALS

Synthetic Data and Multimedia

December 13, 13:00 - 15:00

Image and video processing has seen a rapid growth in the last decade, with remarkable improvements made possible thanks to the availability of ever-increasing computing power as well as deep learning-based frameworks that now allow human-like and beyond performances in many applications, including detection, classification, segmentation, to name a few. However, it is to be noted that the development of novel algorithms and solutions is strictly bound to the availability of a relevant amount of data, which must be representative of the task that needs to be addressed. With this respect, the literature has shown a rapid proliferation of datasets, tackling a multitude of problems, from the simplest to the most complex ones. Some of them are largely adopted and are currently recognized as the reference benchmark against which all newly proposed methods need to compete. Still, there is an ever growing demand for data, to which researchers respond with larger and larger datasets, at a huge cost in terms of acquisition, storage, and annotation of images and clips, often facing inconsistencies in annotations. The use of synthetically-generated data can overcome such limitations, as the generation engine can be designed to fulfill an arbitrary number of requirements, all at the same time. The tutorial will take a holistic view on the ongoing research, the relevant issues, and the potential application of using synthetic data for multimedia data processing, as a standalone resource or in combination with real data. In particular, the attention will be focused on the domain of images and videos, where the lack of representative data for specific problem categories has let emerge the possibility of relying on machine-generated contents.

Nicola Conci
University of Trento

Niccolò Bisagno
University of Trento

Human-centric Visual Understanding

December 13, 15:00 - 17:00

Human-centric visual understanding is one of the fundamental problems of computer vision and multimedia understanding. With the development of deep learning and multi-modalities analysis techniques, researchers have strived to push the limits of human-centric visual understanding in a wide variety of applications, such as intelligent surveillance, retailing, fashion design, and services. This tutorial will present recent advances under the umbrella of human-centric visual understanding, ranging from the fundamental problems of gait recognition, monocular real-time 3D human recovery, human action understanding, and motion prediction, human action analysis in surveillance videos, finally to multimedia event analysis and understanding in complex scenarios and industrial applications. We will discuss the key problems, common formulations, existing methodologies, real industrial applications, future directions, etc. The views of our tutorial not only come from the research field but also combine the real-world requirements and experiences in the industrial community. Therefore, this tutorial will inspire audiences from the research and industrial community, and facilitate research in computer vision and multimedia for human behavior analysis and human-centric analysis modeling. We held a tutorial on the same topic for the first time in ACM Multimedia Asia 2019, so we mainly introduce our work from 2019 to 2022 in this tutorial.
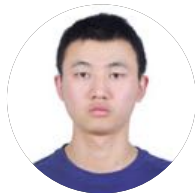
Kun Liu
JD Logistics
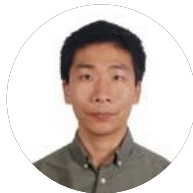
Hongsong Wang
Southeast University

Xinchen Liu
JD Explore Academy

Qian Bao
JD Explore Academy

Cheng Zhang
Carnegie Mellon University

Wu Liu
JD Explore Academy

WORKSHOP

Multimedia Understanding with Pre-trained Models

December 13, 09:45 am – 11:45 am

Multi-modal understanding plays a crucial role in enabling the machine to perceive the physical world with multiple sensor cues as humans. Recently, large-scale pre-trained models (PTMs) has become a research hotspot in the field of artificial intelligence. Existing techniques follow the self-supervised learning paradigm achieve great success on the uni-modal scenes, such as computer vision (CV) and natural language process (NLP). The recent advances in large-scale pre-trained models inspire the researchers to explore more and deeper pre-training techniques for the multi-modal understanding problem. In this workshop, we aim to bring together researchers from the field of multimedia to discuss recent research and future directions on pre-trained models with self-supervised learning for multimedia understanding.

In recent years, we have witnessed the great success of pre-trained models (PTM) in natural language processing (NLP), such as GPT3, BERT, Roberta, DEBERTA, etc. It motivates the researchers in the multimedia community to leverage the idea of PTM to address multi-modal tasks. The scope of this workshop is focused on pre-trained models with self-supervised learning for multimedia understanding. The potential topics include architecture design for multi-modal PTM, pre-text task design for self-supervised learning, multi-modal data modeling, efficiency enhancing for PTM, interpretability of PTM, etc.

PROGRAM

December 13

9:15-	Registration
9:45-11:45	REMOTE Workshop: Multimedia Understanding with Pre-trained Models
13:00-15:00	ONSITE Tutorial 1: Synthetic Data and Multimedia
15:00-17:00	REMOTE Tutorial 2: Human-centric Visual Understanding

December 14

9:45-10:00 Opening

10:00-11:40 Oral 1-1 Award session
Ichijo Hall
Chair: Wei-Ta Chu (National Cheng Kung University)

89 ONSITE TFM a Dataset for Detection and Recognition of Masked Faces in the Wild

Gibran Benitez-Garcia (The University of Electro-Communications)*; Miguel Jimenez-Martinez (INSTITUTO POLITECNICO NACIONAL); Jesus Olivares-Mercado (INSTITUTO POLITECNICO NACIONAL); Hiroki Takahashi (the University of Electro-Communications)

93 REMOTE Deep Image and Kernel Prior Learning for Blind Super-Resolution

Kazuhiro Yamawaki (Yamaguchi University); Xian-Hua Han (Yamaguchi University)*

42 REMOTE Asymmetric Label Propagation for Video Object Segmentation

Zhen Chen (Peking University); Ming Yang (Horizon Robotics); Shiliang Zhang (Peking University)*

39 REMOTE Informative Sample-Aware Proxy for Deep Metric Learning

Aoyu Li (Tokyo Institute of Technology)*; Ikuro Sato (Tokyo Institute of Technology / Denso IT Laboratory); Kohta Ishikawa (Denso IT Laboratory, Inc.); Rei Kawakami (Tokyo Institute of Technology); Rio Yokota (Tokyo Institute of Technology)

83 REMOTE Federated Knowledge Transfer for Heterogeneous Visual Models

Wenzhe Li (Tsinghua University)*; Zirui Zhu (Tsinghua University); Tianchi Huang (Tsinghua University); Lifeng Sun (Tsinghua University); Chun Yuan (Graduate school at ShenZhen, Tsinghua university)

13:00-14:00 Keynote 1
Ichijo Hall
Chair: Keiji Yanai (The University of Electro-Communications)

ONSITE Machine Learning for Creative Workflow

Kota Yamaguchi (CyberAgent, Inc.)

14:00-14:30 Demo Spotlight
Ichijo Hall
Chair: Yoko Yamakata (The University of Tokyo)

100 ONSITE A Music Loop Sequencer with User-adaptive Music Loop Selection

Yuki Iwamoto (Nihon University)*; Tetsuro Kitahara (Nihon University)

105 ONSITE Action Detection System based on Pose Information

Ryo Kawai (NEC Corporation)*; Noboru Yoshida (NEC Corporation); Jianquan Liu (NEC Corporation)

106 ONSITE DeepHair: a DeepFake-based Hairstyle Preview System

Yu-Hsuan Lo (Taipei National University of the Arts); Shih-Wei Sun (Taipei National University of the Arts)*

107 ONSITE Emotional Talking Faces: Making Videos More Expressive and Realistic

Sahil Goyal (IIT Roorkee)*; Shagun Uppal (IIIT-Delhi); Sarthak Bhagat (IIIT-Delhi); Dhroov Goel (IIITD);

Sakshat Mali (Indraprastha Institute of Information Technology, Delhi); Yi Yu (NII); Yifang Yin (A*STAR); Rajiv Ratn Shah (IIIT Delhi)

109 `ONSITE` **FoodLog Athl: Multimedia Food Recording Platform for Dietary Guidance and Food Monitoring**

Kei Nakamoto (The University of Tokyo)*; Kohei Kumazawa (The University of Tokyo); Hiroaki KARASAWA (Hongo Software Development); Sosuke Amano (foo.log Inc.); Yoko Yamakata (University of Tokyo, Japan); Kiyoharu Aizawa (The University of Tokyo)

110 `ONSITE` **Rubber material retrieval system using electron microscope images for rubber material development**

Rintaro Yanagi (Hokkaido University)*; Ren Togo (Hokkaido University); Takahiro Ogawa (Hokkaido University); Miki Haseyama (Hokkaido University)

104 `ONSITE` **JamSketch Deep α: A CNN-based Improvisation System in Accordance with User's Melodic Outline Drawing**

Tetsuro Kitahara (Nihon University)*; Akio Yonamine (Nihon University)

108 `REMOTE` **GSTH266enc: A GStreamer plugin for VVC encoder**

Advaiit Rajjvaed (InterDigital Inc.); Saurabh Puri (InterDigital Inc.)*; Gurdeep Bhullar (InterDigital Inc.); Gaelle Martin-Cocher (InterDigital Inc.)

101 `REMOTE` **Intelligent Video Surveillance Platform Based On FFmpeg And Yolov5**

Chuanxu Jiang (HoHai university); Yanfang Wang (Hohai University); Qian Huang (Hohai University)*; Yiming Wang (HoHai university); Yuhan Dai (Hohai University)

14:30 – 15:45
Lobby

Poster+Demo

100 `ONSITE` `DEMO` **A Music Loop Sequencer with User-adaptive Music Loop Selection**

Yuki Iwamoto (Nihon University)*; Tetsuro Kitahara (Nihon University)

105 `ONSITE` `DEMO` **Action Detection System based on Pose Information**

Ryo Kawai (NEC Corporation)*; Noboru Yoshida (NEC Corporation); Jianquan Liu (NEC Corporation)

106 `ONSITE` `DEMO` **DeepHair: a DeepFake-based Hairstyle Preview System**

Yu-Hsuan Lo (Taipei National University of the Arts); Shih-Wei Sun (Taipei National University of the Arts)*

107 `ONSITE` `DEMO` **Emotional Talking Faces: Making Videos More Expressive and Realistic**

Sahil Goyal (IIT Roorkee)*; Shagun Uppal (IIIT-Delhi); Sarthak Bhagat (IIIT-Delhi); Dhroov Goel (IIITD); Sakshat Mali (Indraprastha Institute of Information Technology, Delhi); Yi Yu (NII); Yifang Yin (A*STAR); Rajiv Ratn Shah (IIIT Delhi)

109 `ONSITE` `DEMO` **FoodLog Athl: Multimedia Food Recording Platform for Dietary Guidance and Food Monitoring**

Kei Nakamoto (The University of Tokyo)*; Kohei Kumazawa (The University of Tokyo); Hiroaki KARASAWA (Hongo Software Development); Sosuke Amano (foo.log Inc.); Yoko Yamakata (University of Tokyo, Japan); Kiyoharu Aizawa (The University of Tokyo)

110 `ONSITE` `DEMO` **Rubber Material Retrieval System using Electron Microscope Images for Rubber Material Development**

Rintaro Yanagi (Hokkaido University)*; Ren Togo (Hokkaido University); Takahiro Ogawa (Hokkaido University); Miki Haseyama (Hokkaido University)

104 `ONSITE` `DEMO` **JamSketch Deep α: A CNN-based Improvisation System in Accordance with User's Melodic Outline Drawing**

Tetsuro Kitahara (Nihon University)*; Akio Yonamine (Nihon University)

108 `REMOTE` `DEMO` **GSTH266enc: A GStreamer plugin for VVC encoder**

Advaiit Rajjvaed (InterDigital Inc.); Saurabh Puri (InterDigital Inc.)*; Gurdeep Bhullar (InterDigital Inc.); Gaelle Martin-Cocher (InterDigital Inc.)

101 `REMOTE` `DEMO` **Intelligent Video Surveillance Platform Based On FFmpeg And Yolov5**

Chuanxu Jiang (HoHai university); Yanfang Wang (Hohai University); Qian Huang (Hohai University)*; Yiming Wang (HoHai university); Yuhan Dai (Hohai University)

89 `ONSITE` `ORAL 1-1` **TFM a Dataset for Detection and Recognition of Masked Faces in the Wild**

Gibran Benitez-Garcia (The University of Electro-Communications)*; Miguel Jimenez-Martinez (INSTITUTO POLITECNICO NACIONAL); Jesus Olivares-Mercado (INSTITUTO POLITECNICO NACIONAL); Hiroki Takahashi (the University of Electro-Communications)

93 `REMOTE` `ORAL 1-1` **Deep Image and Kernel Prior Learning for Blind Super-Resolution**

Kazuhiro Yamawaki (Yamaguchi University); Xian-Hua Han (Yamaguchi University)*

42 `REMOTE` `ORAL 1-1` **Asymmetric Label Propagation for Video Object Segmentation**

Zhen Chen (Peking University); Ming Yang (Horizon Robotics); Shiliang Zhang (Peking University)*

39 `REMOTE` `ORAL 1-1` **Informative Sample-Aware Proxy for Deep Metric Learning**

Aoyu Li (Tokyo Institute of Technology)*; Ikuro Sato (Tokyo Institute of Technology / Denso IT Laboratory); Kohta Ishikawa (Denso IT Laboratory, Inc.); Rei Kawakami (Tokyo Institute of Technology); Rio Yokota (Tokyo Institute of Technology)

83 `REMOTE` `ORAL 1-1` **Federated Knowledge Transfer for Heterogeneous Visual Models**

Wenzhe Li (Tsinghua University)*; Zirui Zhu (Tsinghua University); Tianchi Huang (Tsinghua University); Lifeng Sun (Tsinghua University); Chun Yuan (Graduate school at ShenZhen, Tsinghua university)

1 `REMOTE` `ORAL 1-2` **An End-to-End Scene Text Detector with Dynamic Attention**

Jingyu Lin (厦门大学); Yan Yan (Xiamen University); Hanzi Wang (Xiamen University)*

50 `REMOTE` `ORAL 1-2` **Self-Attentive CLIP Hashing for Unsupervised Cross-Modal Retrieval**

Heng Yu (Nanjing University of Science and Technology); Shuyan Ding (Nanjing University of Science and Technology)*; Lunbo Li (Nanjing University of Science and Technology); Jiexin Wu (Nanjing university of Science and Technology)

66 `ONSITE` `ORAL 1-2` **Affective Embedding Framework with Semantic Representations from Tweets for Zero-shot Visual Sentiment Prediction**

Yingrui Ye (Hokkaido University)*; Yuya Moroto (Hokkaido University); Keisuke Maeda (Hokkaido University); Takahiro Ogawa (Hokkaido University); Miki Haseyama (Hokkaido University)

16 `REMOTE` `ORAL 1-2` **SPEAKER VGG CCT: Cross-corpus Speech Emotion Recognition with Speaker Embedding and Vision**

Alessandro Arezzo (University of Florence, Italy); Stefano Berretti (University of Florence, Italy)*

6 `ONSITE` `ORAL 3-2` **Robust Learning with Adversarial Perturbations and Label Noise: A Two-Pronged Defense Approach**

Peng-Fei Zhang (University of Queensland)*; Zi Helen Huang (University of Queensland); Xin Luo (Shandong University); Pengfei Zhao (Shandong University)

70 `ONSITE` `ORAL 3-2` **Enhancing the Robustness of Deep Learning Based Fingerprinting to Improve Deepfake Attribution**

Chieh-Yin Liao (National Taiwan University)*; Chen-hsiu Huang (National Taiwan University); Jun-Cheng Chen (Academia Sinica); Ja-Ling Wu (NTU)

65 `ONSITE` `ORAL 3-2` **Disentangled Image Attribute Editing in Latent Space via Mask-based Retention Loss**

Shunya Ohaga (Hokkaido University)*; Ren Togo (Hokkaido University); Takahiro Ogawa (Hokkaido

University); Miki Haseyama (Hokkaido University)

36 ONSITE ORAL 3-2 ObjectMix: Data Augmentation by Copy-Pasting Objects in Videos for Action Recognition

Jun Kimata (Nagoya Institute of Technology); Tomoya Nitta (Nagoya Insitute of Technology); Toru Tamaki (Nagoya Institute of Technology)*

20 REMOTE ORAL 3-2 CMR3D: Contextualized Multi-Stage Refinement for 3D Object Detection

Dhanalaxmi Gaddam (Mohammed Bin Zayed University of Artificial Intelligence)*; Jean Lahoud (MBZUAI); Fahad Shahbaz Khan (MBZUAI); Rao Muhammad Anwer (MBZUAI/AALTO); Hisham Cholakkal (MBZUAI)

16:00 – 17:00 Oral 1-2 Text, Speech, and Vision
Ichijo Hall   Chair: Rajiv Ratn Shah (IIT Delhi)

66 ONSITE Affective Embedding Framework with Semantic Representations from Tweets for Zero-shot Visual Sentiment Prediction

Yingrui Ye (Hokkaido University)*; Yuya Moroto (Hokkaido University); Keisuke Maeda (Hokkaido University); Takahiro Ogawa (Hokkaido University); Miki Haseyama (Hokkaido University)

16 REMOTE SPEAKER VGG CCT: Cross-corpus Speech Emotion Recognition with Speaker Embedding and Vision

Alessandro Arezzo (University of Florence, Italy); Stefano Berretti (University of Florence, Italy)*

50 REMOTE Self-Attentive CLIP Hashing for Unsupervised Cross-Modal Retrieval

Heng Yu (Nanjing University of Science and Technology); Shuyan Ding (Nanjing University of Science and Technology)*; Lunbo Li (Nanjing University of Science and Technology); Jiexin Wu (Nanjing university of Science and Technology)

1 REMOTE An End-to-End Scene Text Detector with Dynamic Attention

Jingyu Lin (厦门大学); Yan Yan (Xiamen University); Hanzi Wang (Xiamen University)*

December 15

10:00 - 11:30 Oral 2-1 Video Compression, Broadcasting, and Analysis
Ichijo Hall   Chair: Jianquan Liu (NEC)

11 ONSITE Human-Avatar Interaction in Metaverse: Framework for Full-body Interaction

Kit Yung Lam (Hong Kong University of Science and Technology)*; Liang Yang (Hong Kong University of Science and Technology); Ahmad ALHILAL (The Hong Kong University of Science and Technology); Lik Hang Lee (KAIST); Gareth Tyson (Hong Kong University of Science and Technology (GZ)); Pan HUI (The Hong Kong University of Science and Technology)

49 ONSITE Parallel Queries for Human-Object Interaction Detection

Junwen Chen (The University of Electro-Communications)*; Keiji Yanai (Univ. Electro-Comm., Tokyo)

95 REMOTE Sequential Frame-Interpolation and DCT-based Video Compression Framework

Yeganeh Jalalpour (Portland State University)*; Wu-chi Feng (Portland State University); Feng Liu (Portland State University)

25 REMOTE 360BroadView: Viewer Management for Viewport Prediction in 360-Degree Video Live Broadcast

Qian Zhou (University of Illinois at Urbana-Champaign)*; Zhe Yang (University of Illinois at Urbana-Champaign); Hongpeng Guo (University of Illinois at Urbana Champaign); Beitong Tian (University of Illinois at Urbana Champaign); Klara Nahrstedt (University of Illinois at Urbana-Champaign)

72 REMOTE Two-Layer Learning-based P-Frame Coding with Super-Resolution and

Content-Adaptive Conditional ANF

David Alexandre (National Yang Ming Chiao Tung University)*; Hsueh-Ming Hang (National Yang Ming Chiao Tung University); Wen-Hsiao Peng (National Yang Ming Chiao Tung University)

74 REMOTE Learned Bi-Directional Motion Prediction for Video Compression

Yunhui Shi (Beijing University of Technology); Shaopei An (Beijing University of Technology)*; Jin Wang (Beijing University of Technology); Baocai Yin (Beijing University of Technology)

13:00 - 14:00
Ichijo Hall

Keynote 2

Chair: Kiyoharu Aizawa (The University of Tokyo)

ONSITE Connecting the Dots: Digital Humanities and Historical Big Data Research for Japanese Culture

Asanobu Kitamoto (NII)

14:00 - 15:00
Ichijo Hall

Short Spotlight

Chair: Toshihiko Yamasaki (The University of Tokyo)

9 ONSITE A Multimodal Sensor Fusion Framework Robust to Missing Modalities for Person Recognition

Vijay John (RIKEN)*; Yasutomo Kawanishi (RIKEN)

34 ONSITE SLGAN: Style- and Latent-guided Generative Adversarial Network for Desirable Makeup Transfer and Removal

Daichi Horita (The University of Tokyo)*; Kiyoharu Aizawa (The University of Tokyo)

64 ONSITE Popularity-aware Graph Social Recommendation for Fully Non-Interaction Users

Nozomu Onodera (Hokkaido University)*; Keisuke Maeda (Hokkaido University); Takahiro Ogawa (Hokkaido University); Miki Haseyama (Hokkaido University)

28 ONSITE Multimodal Fusion with Cross-Modal Attention for Action Recognition in Still Images

Jia-Hua Tsai (Department of Computer Science and Information Engineering, National Cheng-Kung University); Wei-Ta Chu (National Cheng Kung University)*

45 ONSITE Zero-shot Font Style Transfer with a Differentiable Renderer

Kota Izumi (The University of Electro-Communications); Keiji Yanai (Univ. Electro-Comm., Tokyo)*

98 ONSITE Wearable Camera Based Food Logging System

Kenshiro Sato (The University of Tokyo)*; Yoko Yamakata (University of Tokyo, Japan); Sosuke Amano (foo. log Inc.); Kiyoharu Aizawa (The University of Tokyo)

90 ONSITE Graph Neural Network Based Living Comfort Prediction Using Real Estate Floor Plan Images

Ryota Kitabayashi (The University of Tokyo)*; Taro Narahara (njit); Toshihiko Yamasaki (The University of Tokyo)

48 ONSITE Wider or Deeper Neural Network Architecture for Acoustic Scene Classification with Mismatched Recording Devices

Lam Pham (Austrian Institute of Technology)*; Khoa Tran (Da Nang University); Dat Ngo (University of Essex); Hieu Tang (FPT University); Son Phan (Earable AI); Alexander Schindler (Austrian Institute of Technology)

24 REMOTE A Reality Check of Positioning in Multiuser Mobile Augmented Reality: Measurement and Analysis

Na Wang (George Mason University)*; Haoliang Wang (Adobe Research); Stefano Petrangeli (Adobe); Viswanathan (Vishy) Swaminathan (Adobe); Fei Li (George Mason University); Songqing Chen (George Mason University)

59 REMOTE Towards High Performance One-Stage Human Pose Estimation

Ling Li (Nanjing University of Science and Technology); Lin Zhao (Nanjing University of Science and Technology)*; Linhao Xu (Nanjing University of Science and Technology); Jie Xu (Nanjing University of Science and Technology)

85 REMOTE Singing Voice Detection via Similarity-based Semi-supervised Learning Method

Xi Chen (Fudan University)*; Yongwei Gao (Shanghai University of International Business and Economics); Wei Li (Fudan University)

15:15 - 17:00   Poster+Demo

Lobby

100 ONSITE DEMO A Music Loop Sequencer with User-adaptive Music Loop Selection

Yuki Iwamoto (Nihon University)*; Tetsuro Kitahara (Nihon University)

105 ONSITE DEMO Action Detection System based on Pose Information

Ryo Kawai (NEC Corporation)*; Noboru Yoshida (NEC Corporation); Jianquan Liu (NEC Corporation)

106 ONSITE DEMO DeepHair: a DeepFake-based Hairstyle Preview System

Yu-Hsuan Lo (Taipei National University of the Arts); Shih-Wei Sun (Taipei National University of the Arts)*

107 ONSITE DEMO Emotional Talking Faces: Making Videos More Expressive and Realistic

Sahil Goyal (IIT Roorkee)*; Shagun Uppal (IIIT-Delhi); Sarthak Bhagat (IIIT-Delhi); Dhroov Goel (IIITD); Sakshat Mali (Indraprastha Institute of Information Technology, Delhi); Yi Yu (NII); Yifang Yin (A*STAR); Rajiv Ratn Shah (IIIT Delhi)

109 ONSITE DEMO FoodLog Athl: Multimedia Food Recording Platform for Dietary Guidance and Food Monitoring

Kei Nakamoto (The University of Tokyo)*; Kohei Kumazawa (The University of Tokyo); Hiroaki KARASAWA (Hongo Software Development); Sosuke Amano (foo.log Inc.); Yoko Yamakata (University of Tokyo, Japan); Kiyoharu Aizawa (The University of Tokyo)

110 ONSITE DEMO Rubber material retrieval system using electron microscope images for rubber material development

Rintaro Yanagi (Hokkaido University)*; Ren Togo (Hokkaido University); Takahiro Ogawa (Hokkaido University); Miki Haseyama (Hokkaido University)

104 ONSITE DEMO JamSketch Deep α: A CNN-based Improvisation System in Accordance with User's Melodic Outline Drawing

Tetsuro Kitahara (Nihon University)*; Akio Yonamine (Nihon University)

108 REMOTE DEMO GSTH266enc: A GStreamer plugin for VVC encoder

Saurabh Puri (InterDigital Inc.)*; Advaiit Rajjvaed (InterDigital Inc.); Gurdeep Bhullar (InterDigital Inc.); Gaelle Martin-Cocher (InterDigital Inc.)

101 REMOTE DEMO Intelligent Video Surveillance Platform Based On FFmpeg And Yolov5

Chuanxu Jiang (HoHai university); Qian Huang (Hohai University)*; Yiming Wang (HoHai university); Yuhan Dai (Hohai University); Yanfang Wang (Hohai University)

9 ONSITE SHORT A Multimodal Sensor Fusion Framework Robust to Missing Modalities for Person Recognition

Vijay John (RIKEN)*; Yasutomo Kawanishi (RIKEN)

34 ONSITE SHORT SLGAN: Style- and Latent-guided Generative Adversarial Network for Desirable Makeup Transfer and Removal

Daichi Horita (The University of Tokyo)*; Kiyoharu Aizawa (The University of Tokyo)

64 ONSITE SHORT Popularity-aware Graph Social Recommendation for Fully Non-Interaction Users

Nozomu Onodera (Hokkaido University)*; Keisuke Maeda (Hokkaido University); Takahiro Ogawa (Hokkaido University); Miki Haseyama (Hokkaido University)

28 ONSITE SHORT Multimodal Fusion with Cross-Modal Attention for Action Recognition in Still Images

Jia-Hua Tsai (Department of Computer Science and Information Engineering, National Cheng-Kung University); Wei-Ta Chu (National Cheng Kung University)*

45 ONSITE SHORT Zero-shot Font Style Transfer with a Differentiable Renderer

Kota Izumi (The University of Electro-Communications); Keiji Yanai (Univ. Electro-Comm., Tokyo)*

98 ONSITE SHORT Wearable Camera Based Food Logging System

Kenshiro Sato (The University of Tokyo)*; Yoko Yamakata (University of Tokyo, Japan); Sosuke Amano (foo. log Inc.); Kiyoharu Aizawa (The University of Tokyo)

90 ONSITE SHORT Graph Neural Network Based Living Comfort Prediction Using Real Estate Floor Plan Images

Ryota Kitabayashi (The University of Tokyo)*; Taro Narahara (njit); Toshihiko Yamasaki (The University of Tokyo)

48 ONSITE SHORT Wider or Deeper Neural Network Architecture for Acoustic Scene Classification with Mismatched Recording Devices

Lam Pham (Austrian Institute of Technology)*; Khoa Tran (Da Nang University); Dat Ngo (University of Essex); Hieu Tang (FPT University); Son Phan (Earable AI); Alexander Schindler (Austrian Institute of Technology)

24 REMOTE SHORT A Reality Check of Positioning in Multiuser Mobile Augmented Reality: Measurement and Analysis

Na Wang (George Mason University)*; Haoliang Wang (Adobe Research); Stefano Petrangeli (Adobe); Viswanathan (Vishy) Swaminathan (Adobe); Fei Li (George Mason University); Songqing Chen (George Mason University)

59 REMOTE SHORT Towards High Performance One-Stage Human Pose Estimation

Ling Li (Nanjing University of Science and Technology)*; Lin Zhao (Nanjing University of Science and Technology); Linhao Xu (Nanjing University of Science and Technology); Jie Xu (Nanjing University of Science and Technology)

85 REMOTE SHORT Singing Voice Detection via Similarity-based Semi-supervised Learning Method

Xi Chen (Fudan University)*

11 ONSITE ORAL 2-1 Human-Avatar Interaction in Metaverse: Framework for Full-body Interaction

Kit Yung Lam (Hong Kong University of Science and Technology)*; Ahmad ALHILAL (The Hong Kong University of Science and Technology); Liang Yang (Hong Kong University of Science and Technology); Lik Hang Lee (The University of Oulu); Gareth Tyson (Queen Mary University of London); Pan HUI (The Hong Kong University of Science and Technology)

49 ONSITE ORAL 2-1 Parallel Queries for Human-Object Interaction Detection

Junwen Chen (The University of Electro-Communications)*; Keiji Yanai (Univ. Electro-Comm., Tokyo)

95 REMOTE ORAL 2-1 Sequential Frame-Interpolation and DCT-based Video Compression Framework

Yeganeh Jalalpour (Portland State University)*; Wu-chi Feng (Portland State University); Feng Liu (Portland State University)

25 REMOTE ORAL 2-1 360BroadView: Viewer Management for Viewport Prediction in

360-Degree Video Live Broadcast

Qian Zhou (University of Illinois at Urbana-Champaign)*; Zhe Yang (University of Illinois at Urbana-Champaign); Hongpeng Guo (University of Illinois at Urbana Champaign); Beitong Tian (University of Illinois at Urbana Champaign); Klara Nahrstedt (University of Illinois at Urbana-Champaign)

72 REMOTE ORAL 2-1 **Two-Layer Learning-based P-Frame Coding with Super-Resolution and Content-Adaptive Conditional AN**

David Alexandre (National Yang Ming Chiao Tung University)*; Hsueh-Ming Hang (National Yang Ming Chiao Tung University); Wen-Hsiao Peng (National Yang Ming Chiao Tung University)

74 REMOTE ORAL 2-1 **Learned Bi-Directional Motion Prediction for Video Compression**

Yunhui Shi (Beijing University of Technology)*; Shaopei An (Beijing University of Technology); Jin Wang (Beijing University of Technology); Baocai Yin (Beijing University of Technology)

55 ONSITE ORAL 3-1 **Deep Enhancement-Object Features Fusion for Low-light Object Detection**

Wan Teng Lim (Multimedia University); Kelvin Ang (Multimedia University); Yuen Peng Loh (Multimedia University)*

7 REMOTE ORAL 3-1 **Image Compression for Machines Using Boundary-Enhanced Saliency**

Yuanyuan Xu (Hohai University)*; Haolun Lan (Hohai University)

30 REMOTE ORAL 3-1 **Deep Weighted Guided Upsampling Network for Depth of Field Image Upsampling**

Lanling Zeng (Jiangsu University)*; Lianxiong Wu (Jiangsu University); Yang Yang (Jiangsu University); Xiang-Jun Shen (Jiangsu University); Yongzhao Zhan (UJS)

57 REMOTE ORAL 3-1 **Multispectral Image Denoising Via Structural Tensor Sparsity Promoting Model**

longlu huang (Beijing University of Technology); Na Qi (Beijing University of Technology)*; Qing Zhu (Beijing University of Technology)

53 REMOTE ORAL 3-1 **Multi-scale Channel Transformer Network for Single Image Deraining**

Yuto Namba (Yamaguchi University)*; Xian-Hua Han (Yamaguchi University)

67 REMOTE ORAL 3-1 **Remote sensing image colorization based on Joint Stream Deep Convolutional Generative Adversarial Networks**

Jingyu J Wang (Ocean University of China); Jie Nie (Ocean University of China)*; Hao Chen (Ocean University of China); Huaxin Xie (Ocean University of China); Chengyu Zheng (Ocean University of China); Min Ye (Ocean university of China); Zhiqiang Wei (Ocean University of China)

86 ONSITE ORAL 3-2 **On the Robustness of 3D Object Detectors**

Fatima A Albreiki (MBZUAI); Sultan Abughazal (MBZUAI)*; Jean Lahoud (MBZUAI); Rao Muhammad Anwer (MBZUAI/AALTO); Hisham Cholakkal (MBZUAI); Fahad Shahbaz Khan (MBZUAI)

December 16

10:00 - 11:30   Oral 3-1 Low-level Vision and Image Processing
Ichijo Hall   Chair: Rei Kawakami (Tokyo Institute of Technology)

55 ONSITE **Deep Enhancement-Object Features Fusion for Low-light Object Detection**

Wan Teng Lim (Multimedia University); Kelvin Ang (Multimedia University); Yuen Peng Loh (Multimedia University)*

7 `REMOTE` **Image Compression for Machines Using Boundary-Enhanced Saliency**

Yuanyuan Xu (Hohai University)*; Haolun Lan (Hohai University)

30 `REMOTE` **Deep Weighted Guided Upsampling Network for Depth of Field Image Upsampling**

Lanling Zeng (Jiangsu University)*; Lianxiong Wu (Jiangsu University); Yang Yang (Jiangsu University); Xiang-Jun Shen (Jiangsu University); Yongzhao Zhan (UJS)

57 `REMOTE` **Multispectral Image Denoising Via Structural Tensor Sparsity Promoting Model**

Longlu Huang (Beijing University Of Technology); Na Qi (Beijing University of Technology)*; Qing Zhu (Beijing University of Technology)

53 `REMOTE` **Multi-scale Channel Transformer Network for Single Image Deraining**

Yuto Namba (Yamaguchi University)*; Xian-Hua Han (Yamaguchi University)

67 `REMOTE` **Remote sensing image colorization based on Joint Stream Deep Convolutional Generative Adversarial Networks**

Jingyu J Wang (Ocean University of China); Jie Nie (Ocean University of China)*; Hao Chen (Ocean University of China); Huaxin Xie (Ocean University of China); chengyu zheng (Ocean University of China); Min Ye (Ocean university of China); Zhiqiang Wei (Ocean University of China)

13:00 - 14:30
Ichijo Hall

Oral 3-2 Robustness, Data Augmentation and Disentangling
Chair: Yasushi Makihara (Osaka University)

86 `ONSITE` **On the Robustness of 3D Object Detectors**

Fatima A Albreiki (MBZUAI); Sultan Abughazal (MBZUAI)*; Jean Lahoud (MBZUAI); Rao Anwer (MBZUAI); Hisham Cholakkal (MBZUAI); Fahad Shahbaz Khan (MBZUAI)

6 `ONSITE` **Robust Learning with Adversarial Perturbations and Label Noise: A Two-Pronged Defense Approach**

Peng-Fei Zhang (University of Queensland)*; Zi Helen Huang (University of Queensland); Xin Luo (Shandong University); Pengfei Zhao (Shandong University)

70 `ONSITE` **Enhancing the Robustness of Deep Learning Based Fingerprinting to Improve Deepfake Attribution**

Chieh-Yin Liao (National Taiwan University)*; Chen-hsiu Huang (National Taiwan University); Jun-Cheng Chen (Academia Sinica); Ja-Ling Wu (NTU)

65 `ONSITE` **Disentangled Image Attribute Editing in Latent Space via Mask-based Retention Loss**

Shunya Ohaga (Hokkaido University)*; Ren Togo (Hokkaido University); Takahiro Ogawa (Hokkaido University); Miki Haseyama (Hokkaido University)

36 `ONSITE` **ObjectMix: Data Augmentation by Copy-Pasting Objects in Videos for Action Recognition**

Jun Kimata (Nagoya Institute of Technology); Tomoya Nitta (Nagoya Insitute of Technology); Toru Tamaki (Nagoya Institute of Technology)*

20 `REMOTE` **CMR3D: Contextualized Multi-Stage Refinement for 3D Object Detection**

Dhanalaxmi Gaddam (Mohammed Bin Zayed University of Artificial Intelligence)*; Jean Lahoud (MBZUAI); Fahad Shahbaz Khan (MBZUAI); Rao Muhammad Anwer (MBZUAI/AALTO); Hisham Cholakkal (MBZUAI)
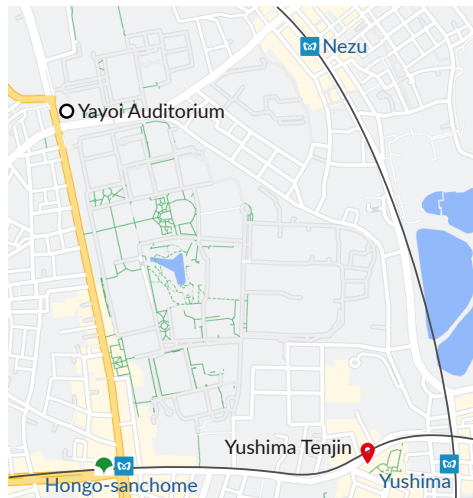
BANQUET

December 15, 6:30pm
Banquet Hall
Yushima Tenjin Shrine

This event is included with the ONSITE registration.

You may enjoy a 20-min walk from the conference venue to the banquet venue or use a subway for quicker access.

📍Google map

Yushima Station
Chiyoda Line
Exit 3
2 min on foot

FINANCIAL SPONSOR

SUBSIDY

公益財団法人
鹿島学術振興財団
THE KAJIMA FOUNDATION

公益財団法人
電気通信普及財団
The Telecommunications Advancement Foundation

公益財団法人
立石科学技術振興財団

SUPPORTER