

## Guest Editorial: Special Issue on “Intelligent-based Image and Vision Computing for Future Video Surveillance”

Xiaoxian Yang\*, Jung Yoon Kim, Walayat Hussain

The video surveillance applications typically require massive devices with real-time computation capability for monitoring, recognition and data analysis. These tasks bring several challenging issues in developing intelligent solutions for handling large-scale and continuous video data. In this special issue, we selected five papers from submissions. A summary of these papers is outlined below.

In the paper entitled “Edge Computing Offloading at Middle-sea Scenario for Maritime Video Surveillances” by Ziyang Gong et al., the authors combine mobile edge computing with abundant node resources, and the network connectivity delay characteristics to handle massive maritime video surveillance information. They establish SSU (single-user single-hop unicast) and multi-user single-hop unicast models (MSUS1 and MSUS2) based on the middle-sea scenario, and divide the optimization problems into two sub-problems. A binary search method is proposed to optimize the transmission power allocation. The methods of OAAS (offloading algorithm based on alternating selection), OAMOAS (multi-objective alternating selection), and OANR (node redistribution offloading algorithm) are proposed to optimize the offloading decision allocation. The experiment shows that their method outperforms in saving delay.

In the paper entitled “Segmentation-Based Decision Networks for Steel Surface Defect Detection” by Zhongqin Bi et al., the authors explore the use of different numbers of labels with various accuracies during training to achieve the maximum detection accuracy with the lowest cost. The proposed method includes the improved segmentation and decision networks. An attention mechanism is integrated into the segmentation subnetwork. The atrous convolutions are used in the segmentation and decision subnetworks. Then, the original loss function is improved. Experiments are carried out on the Severstal Steel Defect dataset. The results show that the detection accuracy is improved by 1% to 2%.

In the paper entitled “A Multi-modal Feature Fusion-based Approach for Mobile Application Classification and Recommendation” by Buqing Cao et al., the authors propose a mobile application classification and recommendation method based on multimodal feature fusion. First, the method extracts the image and description features of the mobile application using the TRedBert model, such as the integration of involution residual network and pre-trained language representation model. Second, these features are fused by using the attention mechanism in the transformer model. Then, the classification results of the mobile applications are based on Softmax. Finally, the method extracts the high-order and low-order embedding features of the mobile app with

FiBiNET (a bi-linear feature interaction) model, to update the mobile app representation and complete the recommendation task. The experiment results demonstrated that the proposed approach outperforms other methods in terms of F1, Accuracy, AUC, and Log-loss.

In the paper entitled “IAMPDNet: Instance-aware and Multi-part Decoupled Network for Joint Detection and Embedding” by Pan Yang et al., the authors point out that intelligent video surveillance methods have been widely investigated to address largescale video data, among which multi-object tracking (MOT) is the most popular method, which aims to track every object appearing in the video for monitoring. For accelerating the inference speed, the method of joint detection and embedding (JDE) has become a new paradigm for MOT. Thus, the authors propose an instance-aware and multi-part decoupled network (IAMPDNet), which can perceive all instances in the environment and extract multi-part features from the instances. IAMPDNet consists of three key modules: a complementary attention module used to perceive all instances in the environment, a feature extraction module used to decouple multi-part features from the instances, and an adaptive aggregation module used to fuse multi-level features of instances. Experiments on MOT benchmarks demonstrate that IAMPDNet achieves higher tracking accuracy and lower identity switches against recent MOT methods.

In the paper entitled “Time-based calibration: A way to ensure that stitched images are captured simultaneously” by Ziwei Song et al., the authors propose a multi-source video frames calibration technique based on external information sources for solving the problems of ghosting and cutting when stitching with different video sources. The proposed method calibrates the video stitching by introducing an information source and calculating the time difference between different devices. Experiments show that the error of the calibrated video stitching is less than 33 ms, which guarantees the quality of the spliced video.

---

\*Corresponding Author: Xiaoxian Yang; E-mail: xxyang@sspu.edu.cn

## Guest Editors



**Xiaoxian Yang**, School of Computer and Information Engineering, Shanghai Polytechnic University, China. (xxyang@sspu.edu.cn)

Xiaoxian Yang received a Ph.D. in Management Science and Engineering from Shanghai University, China in 2017. She is currently an assistant professor at Shanghai Polytechnic University, China. Dr. Yang's research interests include business process management, formal verification, wireless networks, and mobile health. She has published more than 20 papers in academic journals, such as IEEE TITS, IEEE TCSS, ACM TOMM, ACM TOIT, FGCS, MONET, IJSEKE, FGCS, and COMNET. Dr. Yang has participated in organizing international conferences and workshops, such as CollaborateCom 2018-2020, ChinaCom 2019-2020, and Mobicase 2019-2020. She also worked as a guest editor for MONET, WINE, and JIT and served as a reviewer for IEEE TII, IEEE T-ITS, IEEE T-ASE, IEEE JBHI, FGCS, PPNA, Wireless Networks, COMPUTER NETWORK, etc.



**Jung Yoon Kim**, College of Future Industry, Gachon University, South Korea. (kjyoon@gachon.ac.kr).

Jung Yoon Kim received a Ph.D. degree in game engineering from the Graduate School of Advanced Imaging Science, Multimedia & Film at Chung-Ang University in 2013. He is an Associate Professor at the Graduate School of Game Gachon University and the Department of Game Media, College of Future Industry. He is also the center director of the Start-Up Education Center at Gachon University. He was the vice president of KGDA (Korea Game Developer Associations) from 2015 to April 2018. He has served as the Editor-in-Chief for the Korea Computer Game Association since 2016. His research interests include computer gaming, AI, virtual reality technology, and interactive technology



**Walayat Hussain**, Victoria University, Australia. (walayat.hussain@vu.edu.au)

Walayat Hussain received a Ph.D. degree from the University of Technology Sydney, Australia. Currently, he is serving as a Lecturer (Asst. Professor) at Victoria University, Melbourne, Australia. Before joining Victoria University, he worked for six years as a Lecturer and Research Fellow at the FEIT, University of Technology Sydney, Australia. He has served for many years as an Assistant Professor at BUIITEMS University. He has published in various top-ranked ERA-A\*, JCR/SJR Q1 journals, such as The Computer Journal, Info. Systems, Info. Sciences, IJIS, FGCS, IEEE Access, Comput & Ind Eng, MONET, Journal of AIHC, IEEE TETCI, IEEE TSC, IEEE TGCN, IJCS, and WCMC. He is an Associate Editor of IET Communications and has served as a guest editor for various Q1 journals. He has won multiple national and international

research awards. He was the recipient of the Best Paper Award at 3PGCIC 2015 and 2016 in Poland and South Korea, from the Ministry of Higher Education Govt. of Oman, and the FEIT HDR Publication Award from UTS Australia