



## **ICIRA 2024 Special Session Proposal**

**Title of the Proposal: Large model learning and robotics**

**Technical Outline of the Session and Topics:**

With the advent of ChatGPT, using large intelligent models to drive robotic systems become one of the significant trends. Large intelligent models extend the scope of the application of robot systems, improve the level of intelligent service, optimize the quality of service, and give a better service experience for us. However, as the application of the large model is a new direction, there are several challenges that need to be handled, such as visual perception, natural language processing, audio analysis, information fusion, coordination, planning, decision, interaction, etc.

This special issue aims to publish high-quality papers at the intersection domain of large intelligent models and robotic systems. We solicit original, full-length, unpublished contributions in this domain. Potential topics of interests include, but are not limited to:

- 1) Visual perception
- 2) Natural language processing
- 3) Audio analysis
- 4) Multi-modal fusion
- 5) Control and coordination
- 6) Human-Robot interaction
- 7) Planning and decision
- 8) Model compression

**Contact details of the Session Organizers**

### **Prof. Baochang Zhang**

Baochang Zhang is currently a Professor with Beihang University, and an academic advisor of Deep Learning Lab of Baidu Inc., Beijing, China. He received the B.S., M.S., and Ph.D. degrees in computer science from the Harbin Institute of Technology, Harbin, China, in 1999, 2001, and 2006, respectively. From 2006 to 2008, he was a Research Fellow with The Chinese University of Hong Kong, Hong Kong, and Griffith University, Brisbane, QLD, Australia. He was a Senior Post-Doctoral Fellow with the Italy Institute of Technology, Genoa, Italy, from 2014 to 2015. His current research interests include deep learning, pattern recognition, object recognition and tracking, and wavelets.

### **Linlin Yang [mu4yang@gmail.com](mailto:mu4yang@gmail.com)**

Linlin Yang is a lecturer at Communication University of China, Beijing, China. He received a Ph.D. degree from University of Bonn, Bonn, Germany, and the B.Eng. and M.Eng. degrees from Beihang University, Beijing, China. From 2022 to 2023, he was a research fellow at National University of Singapore, Singapore. His research interest is computer vision and machine learning with a special focus on human/hand pose estimation, semi-/self-supervised learning, multi-/cross-modal learning and deep network quantization. He has published more than 20 papers in IEEE/ACM journals and top academic conferences.  
<https://www.mu4yang.com>

### **Prof. Junbiao Pang**

Junbiao Pang received the B.S. and M.S. degrees in computational fluid dynamics and computer science from the Harbin Institute of Technology, Harbin, China, in 2002 and 2004, respectively, and the Ph.D. degree from the Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China, in 2011. He is currently an Associate Professor with the Faculty of Information, Beijing University of Technology, Beijing. He has authored or co-authored approximately 30 academic papers in publications, such as the IEEE Transactions on Image Processing and ACM Multimedia. His research interests include multimedia and machine learning.

### **Dr. Shaohui Lin**

Shaohui Lin is currently a Researcher and a Zijiang Young Scholar in the School of Computer Science and Technology, East China Normal University (ECNU). I received

Ph.D. from Xiamen University. After that, I joined the Department of Computer Science at National University of Singapore as a postdoc researcher. My research specialty is computer vision, machine learning and deep learning, especially compression and speeding-up of large capacity models.

### **Prof. David Doermann**

David Doermann (Fellow, IEEE) is currently a Professor of Empire Innovation with the University at Buffalo (UB), Buffalo, NY, USA, and the Director of the University at Buffalo Artificial Intelligence Institute. Prior to coming to UB, he was a Program Manager with the Defense Advanced Research Projects Agency, where he developed, selected, and oversaw research and transition funding in computer vision, human language technologies, and voice analytics. From 1993 to 2018, he was a Member of the research faculty with the University of Maryland, College Park, MD, USA. He has more than 250 publications in conferences and journals, is a Fellow of the IAPR, has numerous awards, including an honorary doctorate from the University of Oulu, Finland, and is a founding Editor-in-Chief of the International Journal on Document Analysis and Recognition.

-