**Proposal for ICPR 2020 Workshop**

**FGVRID: Fine-Grained Visual Recognition and re-IDentification**

1. **Proposers**

Shiliang Zhang (Peking University), slzhang.jdl@pku.edu.cn

Guorong Li (University of Chinese Academy of Sciences), liguorong@ucas.edu.cn

Weigang Zhang (Harbin Institute of Technology), wgzhang@hit.edu.cn

Qingming Huang (University of Chinese Academy of Sciences), qmhuang@ucas.ac.cn

Nicu Sebe (University of Trento), niculae.sebe@unitn.it

1. **Topics will be covered**

This workshop will bring together researchers from subfields of patter recognition that have seen growing activity in the past few years: *fine-grained visual recognition and person/vehicle re-identification*.

The ubiquitous surveillance cameras are generating huge amount of videos. Automatic video content analysis and recognition are thus desirable for effective utilization of those data. Fine-Grained Visual Recognition and Re-Identification (FGVRID) aims to accurately identify visual objects and match re-appearing targets, *e.g.*, persons and vehicles from a large set of images and videos. It has the potential to offer an unprecedented possibility for intelligent video processing and analysis, as well as to explore the promising applications on public security.

The proposed FGVRID workshop wishes to bring together researchers from fine-grained visual categorization, as well as person/vehicle ReID communities, and to foster discussions and exchange of ideas between them. FGVRID is not a traditional search or classification task due to its goal of accurately identifying visual objects. First, proper detection algorithms should be designed to locate objects and their parts in videos before proceeding to the identification step. Second, the visual appearance of an object is easily affected by many factors like viewpoint changes and camera parameter differences, *etc*. Third, annotating the fine-grained identity or category cues is expensive and time consuming. Finally, to cope with the large-scale data, scalable indexing or feature coding algorithms should be designed to ensure the online recognition efficiency. Aiming to seek novel solutions and possibilities in FGVRID, this workshop will have in-depth discussions on those issues and aims to go beyond toy datasets and small-scale algorithms. Specifically, the covered topics include, but are not limited to:

* Unsupervised, semi-supervised, and transfer learning algorithms
* Robust object detection and tracking in the wild
* Efficient and effective video representations
* Object parsing and layout estimation
* Large-scale indexing, feature coding, and retrieval algorithms
* Fine-grained visual classification
* New problems and datasets for fine-grained visual recognition and re-identification

We envision the workshop to be a process where rewarding discussions take place and outline research directions for years to come. To this end, we want this workshop to begin before its nominal date and to continue after ICPR:

* **Prior to the workshop**, we will collect opinions from the fine-grained visual recognition and person/vehicle re-identification communities on a variety of related topics;
* **At the workshop**, discussions on research questions as well as further research directions will happen between attendees and invited speakers;
* **After the workshop**, special issues on FGVRID in journals like IEEE T-IP and IJCV will be organized. We will also organize further workshops and challenges.

1. **Relevance and viability**

Fine-grained visual recognition and person/vehicle re-identification require highly discriminative and robust visual features, as well as efficient learning algorithms that can be optimized with limited training data. They are closely related to many pattern recognition and computer vision tasks like 1) object detection and tracking, 2) object pattern analysis for representation learning, 3) indexing and retrieval, 4) few-shot learning and unsupervised learning, *etc*. Therefore, the proposed workshop is closely related with the ICPR conference. Recently, an increasing number of works on related topics like fine-grained visual recognition, person re-identification were being published in premier computer vision conferences and journals. Consequently, the relevance between ICPR and this workshop is very high.

The proposers are professors working closely on fine-grained visual categorization and person/vehicle ReID. They will invite experts on this topic as Technical Program Committee to ensure the quality of accepted oral and poster papers. The proposers plan to invite 3-5 world-renowned speakers to share their latest works. The proposers are either experienced in organizing conferences, workshops, or have rich experience on fine-grained visual recognition and person/vehicle ReID research. Consequently, the FGVRID is clearly viable.

1. **Introduction to the proposers**

The proposers are working closely on fine-grained visual categorization, person/vehicle ReID, and related topics. The group led by Prof. Shiliang Zhang in Peking University has developed several efficient fine-grained visual categorization and multi-camera person ReID systems and have many recent works published in journals and conferences like IEEE T-PAMI, T-IP, ECCV, ICCV, and CVPR. Prof. Guorong Li, Prof. Weigang Zhang, and Prof. Qingming Huang have been working on fine-grained visual categorization and object tracking for more than 10 years. Prof. Nicu Sebe is also widely-known for his important contributions on fine-grained visual categorization and efficient machine learning algorithms.

The proposers are from different research communities with diverse backgrounds. Prof. Shiliang Zhang is from Peking University in China. Prof. Guorong Li and Prof. Qingming Huang are from University of Chinese Academy of Sciences. Prof. Weigang Zhang is from Harbin Institute of Technology in China. Prof. Nicu Sebe is from University of Trento in Italy.

The proposers are experienced in organizing workshops and conferences. They have successfully organized two related workshop in CVPR2017 and CVPR 2019, i.e., the 1st and 2nd Workshop on Target Re-Identification and Multi-Target Multi-Camera Tracking. Prof. Nicu Sebe was a General Chair of ICMR 2017, ACM Multimedia 2013, ACM Multimedia 2022, etc.

**Short bios of proposers:**

**Shiliang Zhang** received the Ph.D. degree in Computer Science from the Institute of Computing Technology, Chinese Academy of Sciences, in 2012. He was a Post-Doctoral Scientist with NEC Laboratories America and a Post-Doctoral Research Fellow with The University of Texas at San Antonio. He is currently a tenure-track Assistant Professor with the School of Electronic Engineering and Computer Science, Peking University. He has authored or co-authored over 70 papers in journals and conferences, including IEEE Trans. on PAMI, IEEE Tans. on Image Processing, IEEE Trans. on Multimedia, ACM Multimedia, CVPR, ICCV, and ECCV. His research interests include large-scale image retrieval and computer vision. He was a recipient of the Outstanding Doctoral Dissertation Awards from the Chinese Academy of Sciences and Chinese Computer Federation, the President Scholarship from the Chinese Academy of Sciences, the NEC Laboratories America Spot Recognition Award, and the Microsoft Research Fellowship, etc. He was a recipient of the Top 10% Paper Award at the IEEE MMSP 2011.

**Guorong Li** is an Associate Professor with the University of Chinese Academy of Sciences, Beijing. She received the B.S. degree in computer science from the Renmin University of China, Beijing, China, in 2006, and the Ph.D. degree in computer science from the Graduate University of the Chinese Academy of Sciences, Beijing, in 2012. Her current research interests include object tracking, video object segmentation, crowd counting and UAV video analysis.

**Weigang Zhang** is an Associate Professor in Harbin Institute of Technology, Weihai and is also a postdoctoral researcher at the University of Chinese Academy of Sciences. He received the Bachelor degree in Computer Science and Technology in 2003, the M.S. and Ph.D. degree in Computer Applied Technology in 2005 and 2016, respectively, all from Harbin Institute of Technology, China. His research interests include video analysis, image processing and pattern recognition. He has published more than 60 academic papers and is the recipient of the Best Student Paper Award at IEEE MIPR 2018.

**Qingming Huang** is a chair professor in the University of Chinese Academy of Sciences and an adjunct research professor in the Institute of Computing Technology, Chinese Academy of Sciences. He graduated with a Bachelor degree in Computer Science in 1988 and Ph.D. degree in Computer Engineering in 1994, both from Harbin Institute of Technology, China. His research areas include multimedia computing, image processing, computer vision and pattern recognition He has published more than 400 academic papers in prestigious international journals including IEEE Trans. on PAMI, IEEE Trans. on Image Processing, IEEE Trans. on Multimedia, IEEE Trans. on CSVT, etc, and top-level conferences such as NIPS, ICCV, CVPR, ACM Multimedia, IJCAI, AAAI, VLDB, ICPR, etc. He is the associate editor of IEEE Trans. on CSVT and Acta Automatica Sinica, and the reviewer of various international journals including IEEE Trans. on PAMI, IEEE Trans. on Image Processing, IEEE Trans. on Multimedia, etc. He is a Fellow of IEEE and has served as general chair, program chair, track chair and TPC member for various conferences, including ACM Multimedia, CVPR, ICCV, ICME, ICMR, PCM, BigMM, PSIVT, etc.

**Nicu Sebe** is a Professor at the University of Trento, Italy, leading the research in the areas of multimedia information retrieval and human behavior understanding. He was the General Co- Chair of the IEEE FG Conference 2008 and ACM Multimedia 2013, and the Program Chair of the International Conference on Image and Video Retrieval in 2007 and 2010, ACM Multimedia 2007 and 2011. He is/was the Program Chair of ICCV 2017, ECCV 2016 and ICPR 2020, and a General Chair of ACM ICMR 2017 and ACM Multimedia 2022. He is a fellow of the IAPR.

1. **Program outline**

The workshop will include both invited talks and original submissions selected through a double-blind review process. The proposed length of this workshop is one day. The rough outline and highlights are as follows:

1. 3-5 speakers will be invited to give talks in the workshop. Each speaker will give a 45mins talk on fine-grained visual recognition and person/vehicle ReID or related topics.
2. Apart from the invited talks, there will be oral sessions and poster sessions. We plan to accept submissions not according to their achieved performance on public datasets, but rather encourage people to take on a research question related to the topic of the workshop. Examples of such questions are:
   * How can fine-grained visual recognition and person/vehicle re-identification benefit each other?
   * How much do object/part detections influence performance in fine-grained visual recognition/re-identification?
   * What are the most challenging issues in fine-grained visual recognition/re-identification?
   * How can fine-grained visual recognition/re-identification and Generative Adversarial Networks (GAN) benefit each other?
   * What are the limitations of existing public benchmark datasets?
   * How to tackle the large-scale fine-grained visual recognition/re-identification problem?
3. **Invited speakers**

**The proposers will invite the following speakers (to be confirmed):**

(1) Wei-Shi Zheng, Sun Yat-sen University, zhwshi@mail.sysu.edu.cn

(2) Alberto Del Bimbo, University of Florence, alberto.delbimbo@unifi.it

(3) Rita Cucchiara, University of Modena and Reggio Emilia, rita.cucchiara@unimore.it

(4) Alan Hanjalic, Delft University of Technology, A.hanjalic@utdelft.nl

(5) Cees G.M. Snoek, University of Amsterdam, cgmsnoek@uva.nl

(6) Elisa Ricci, University of Trento, e.ricci@unitn.it

(7) Qi Tian, Huawei Research, q-tian@hotmail.com

1. **Expected participants**

Fine-grained visual recognition/re-identification are important topics for both the academia and industry. It is also an interdisciplinary research topic among Pattern Recognition, Computer Vision, Multimedia, and Machine Learning fields. For example, the advances in object detection, distance metric learning, and deep learning have significantly boosted the performance of person ReID. Meanwhile, the research efforts in person ReID are potential to provide new possibilities for efficient object detection, zero-shot deep learning, scalable indexing and searching, *etc*. Moreover, the FGVRID would be a unique platform allowing world-renowned researchers to share their thoughts and works on related topics. Therefore, this workshop would interest lots of ICPR attendees. The expected number of attendees is 30-50.

The FGVRID has broad coverage on related topics in Pattern Recognition, Computer Vision, Multimedia, and Machine Learning. It is expected to attract researchers working on both this topic and other related fields to submit papers.

1. **Related workshops**

The followings are some recent workshops focusing on Fine-Grained Visual Categorization, person ReID, visual retrieval, and action recognition:

* Workshop on Fine-Grained Visual Categorization in CVPR 2019, CVPR 2018, CVPR 2017, CVPR 2015, CVPR 2013, and CVPR 2011. Organized by Ryan Farrell, et al.
* Workshop on Target Re-Identification and Multi-Target Multi-Camera Tracking in CVPR 2019. Organized by Ergys Ristani, Liang Zheng, Xiatian Zhu, Shiliang Zhang, Jingdong Wang, Shaogang Gong, Qi Tan, Carlo Tomasi, Richard Hartley
* Workshop on Target Re-Identification and Multi-Target Multi-Camera Tracking in CVPR 2017. Organized by Rita Cucchiara, Wen Gao, Shaogang Gong, Thomas S. Huang, Ergys Ristani, Francesco Solera, Qi Tian, Carlo Tomasi, Simone Calderara, Cees G.M. Snoek, Jingdong Wang, Shiliang Zhang
* Visual Surveillance and Re-identification in ECCV 2014. Organized by Shaogang Gong, Steve Maybank, James Orwell, Marco Cristani, Kaiqi Huang, and Shuicheng Yan
* International Workshop on Large Scale Visual Recognition and Retrieval in CVPR 2015. Organized by Samy Bengio, Fei-Fei Li, Olga Russakovsky, Jason Corso, Jia Deng, and Yuanqing Lin
* Large-Scale Video Search and Mining (LSVSM) in CVPR 2015. Organized by Junsong Yuan, Shih-Fu Chang, and John Smith

As can be seen from the above list, there have been many workshops related to fine-grained visual categorization and person ReID. However, fine-grained visual categorization and person/vehicle ReID have achieved significant progress in 2019, *e.g.*, the Rank-1 Acc. on Market1501 has achieved 95+% by designing powerful deep models and global-local features. Moreover, many new problems are being explored, *e.g.*, the applications of Generative Adversarial Networks for data augmentation and transfer, the unsupervised learning strategies. Also, fine-grained visual recognition and person/vehicle ReID are closely related, but are treated as different topics in previous workshops. Therefore, it is necessary to gather researchers in fine grained visual recognition and person/vehicle ReID communities together and share their latest works and ideas.

1. **Equipment requests**

Equipment for presentation (Projector, Speaker) and video recording equipment.