Fig. 1. Overview of Workshop.

3rd Workshop and Challenge on Deepview: Global Multi-Target Visual Surveillance Based on Real-Time Large-Scale Analysis

1 Workshop Title

3rd Workshop and Challenge on Deepview: Global Multi-Target Visual Surveillance Based on Real-Time Large-Scale Analysis

2 Organizers

We have developed various discussions from the past in the following areas: multi-object multi-target tracking [7, 1], person re-identification [9], license plate recognition [4, 3], face recognition [5], abnormal event detection [8, 6], imbalance data processing [2] etc. Based on our fundamental knowledge, we want to further upgrade our knowledge of the topic through cooperation with various researchers.

Moongu Jeon
Professor, Department of EECS - Gwangju Institute of Science and Technology (GIST)
Gwangju, South Korea
mgjeon@gist.ac.kr

Short Bio: Moongu Jeon received the B.S. degree in architectural engineering from Korea University, Seoul, South Korea, in 1988, and the M.S. and Ph.D. degrees in computer science and scientific computation from the University of Minnesota, Minneapolis, MN, USA, in 1999 and 2001, respectively. As the master's degree researcher, he was involved in optimal control problems with the University of California at Santa Barbara, Santa Barbara, CA, USA, from 2001 to 2003, and then moved to the National Research Council of Canada, where he was involved in the sparse representation of high-dimensional data and the image processing, until July 2005. In 2005, he joined the Gwangju Institute of Science and Technology, Gwangju, South Korea, where he is currently a Full Professor with the School of Electrical Engineering and Computer Science. He has served as a editor of Information Sciences. His current research interests include machine learning, computer vision, and artificial intelligence.

Yuseok Bae
Visual Intelligence Research Section, AI Research Laboratory - Electronics and Telecommunications Research Institute (ETRI)
Daejeon, South Korea
baeys@etri.re.kr

Short Bio: Yuseok Bae received the B.S. and M.S. degrees in computer science from Kyungpook National University, Korea, in 1995 and 1997 respectively. He is currently a principal researcher with the Visual Intelligence Research Section at the Electronics and Telecommunications Research Institute (ETRI), where he has been involved in various projects such as distributed computing, home network middleware, IPTV middleware, big data analytics, and visual artificial intelligence since 1997. He received his Ph.D. degree in computer science from Kyungpook National University in 2011. His research interests include distributed computing, computer vision, deep learning, and visual artificial intelligence.

Jinyoung Moon
Visual Intelligence Research Section, AI Research Laboratory - Electronics and Telecommunications Research Institute (ETRI)
Daejeon, South Korea
jymoon@etri.re.kr

Short Bio: Jinyoung Moon received her M.S. degree in Computer Science and Ph.D. in Industrial Systems Engineering from the Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Rep. of Korea, in 2002 and 2018, respectively. Since 2002, she has been working with the Visual Intelligence Research Section, the Artificial Intelligence Research Laboratory, the Electronics and Telecommunications Research Institute (ETRI), Daejeon, Rep. of Korea. Since 2019, she has also been with the ICT department, the University of Science and Technology (UST), where she is currently an Assistant Professor. Her

main research interests include action recognition, human action localization and detection, temporal moment localization, video QA, and representation learning for vision and language.

3 Workshop and Challenge Introduction

With the development of artificial intelligence technology, the demand for smart development of living spaces has increased. Smart cities are one of these demands, including surveillance systems based on large-scale camera networks. This system can provide various services, such as accurate traffic measurements and intelligent surveillance services. These services, which operate in real-time on large camera networks, include many sub-modules like object detection, tracking, re-identification, and human behavior analysis. For constructing complete visual surveillance systems, some issues like illumination changes, dynamic backgrounds, poor data quality, and the lack of high-quality models should be solved. From this workshop, professional speakers will present their research experience and opinions about the future.

Data lack is one of these problems. Real human data or special objects like military weapons are usually hard to gather due to legal and security issues. Also, collecting large datasets from a real environment is very costly. One of the solutions for this problem is synthetic data. Synthetic data is very cheap and collected quickly by the simulation engine. This proposal provides the challenge plan to utilize these synthetic data.

4 Workshop Topics

We embrace the most advanced deep learning systems, meanwhile being open to classical physically grounded models and feature engineering, as well as any well-motivated combination of the two streams. We will invite speakers with professional knowledge and experience about the following topics:

- Robust recognition and detection in the surveillance videos
- Large-scale multi-camera multi-object application (detection, tracking, re-identification, and others)
- 2D/3D pose estimation in the surveillance videos
- One/few-shot learning in the unconstrained surveillance scenarios
- Anomaly detection in the surveillance videos
- Human behavior analysis and recognition in the surveillance videos
- Generation of visual data for surveillance analysis system
- Privacy-preserving visual learning
- Applications and systems for security and safe

5   Challenge Topics

We will organize 1 or 2 challenges. The task of the first challenge is object detection in synthetic aerial imagery. Participants will get training(50k) and testing datasets, including very small military weapons and objects. Also, the background of the provided dataset will show various environments like sea, forest, and mountain. The evaluation metric will be a traditional metric for object detection, mAP.

The task of the second challenge is domain transfer from synthetic data to real data. Participants should collect unpaired real data and transfer synthetic data to real data using various generative models like GAN and diffusion models. We will request short papers(2-3 pages) or regular papers with their problem-solving process to evaluate these submitted methods.
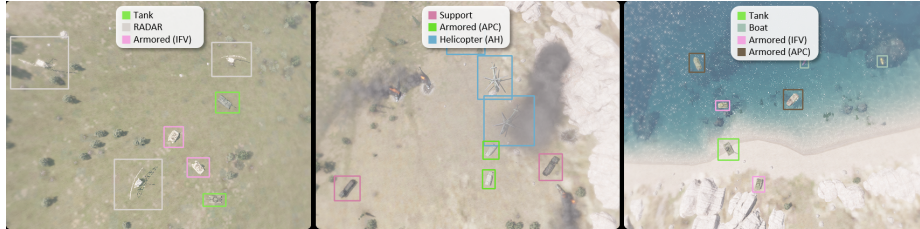

Fig. 1. Overview of dataset.

6   Expected Schedule

A tentative program of the workshop would follow the classic structure of AVSS workshop and challenge programs.

Program	
Opening Remarks	13:00 - 13:15
Invited Talk 1	13:15 - 14:00
Invited Talk 2	14:00 - 14:45
Invited Talk 3	14:45 - 15:30
Coffee Break	15:30 - 16:00
The First Challenge Presentation	16:00 - 16:45
Coffee Break	16:45 - 17:00
The Second Challenge Presentation	17:00 - 17:45
Awards and Closing Remarks	17:45 - 18:00

7 Program Description

We will invite speaker three or more speakers, and they will present for about 45 min. The topic is related to computer vision and intelligent surveillance systems. Challenges will be opened in public challenge management systems like Kaggle. Whole researchers can join these challenges and make teams. We will encourage them to share source code and know-how. It will promote our challenge to make improved results. The top 3 participants of each challenge will present their method for about 15 min. We will provide \$1000 - \$2000 awards for each challenge.

8 Important Dates

Proposal	
Challenge Site Open	August 01, 2023
Challenge Start	September 01, 2023
Challenge End	October 10, 2023
Workshop date	November 6, 2023

Acknowledgements

References

1. Kim, D.Y., Vo, B.N., Vo, B.T., Jeon, M.: A labeled random finite set online multi-object tracker for video data. Pattern Recognition **90**, 377–389 (2019)
2. Kim, Y., Lee, Y., Jeon, M.: Imbalanced image classification with complement cross entropy. Pattern Recognition Letters **151**, 33–40 (2021). https://doi.org/https://doi.org/10.1016/j.patrec.2021.07.017
3. Lee, Y., Jeon, J., Ko, Y., Jeon, M., Pedrycz, W.: License plate detection via information maximization. IEEE Transactions on Intelligent Transportation Systems pp. 1–14 (2021). https://doi.org/10.1109/TITS.2021.3135015
4. Lee, Y., Lee, J., Ahn, H., Jeon, M.: Snider: Single noisy image denoising and rectification for improving license plate recognition. In: Proceedings of the IEEE International Conference on Computer Vision Workshops. pp. 0–0 (2019)
5. Park, S., Yu, J., Jeon, M.: Learning feature representation for face verification. In: 2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS). pp. 1–6. IEEE
6. Yoon, Y., Yu, J., Jeon, M.: Spatio-temporal feature extraction and distance metric learning for unconstrained action recognition. In: 2019 16th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS). pp. 1–8. IEEE (2019)
7. Yoon, Y.C., Kim, D.Y., Song, Y.m., Yoon, K., Jeon, M.: Online multiple pedestrians tracking using deep temporal appearance matching association. Information Sciences **561**, 326–351 (2021)
8. Yu, J., Kim, J.G., Gwak, J., Lee, B.G., Jeon, M.: Abnormal event detection using adversarial predictive coding for motion and appearance. Information Sciences **586**, 59–73 (2022). https://doi.org/https://doi.org/10.1016/j.ins.2021.11.001, https://www.sciencedirect.com/science/article/pii/S0020025521011117
9. Yu, J., Ko, D., Moon, H., Jeon, M.: Deep discriminative representation learning for face verification and person re-identification on unconstrained condition. In: 2018 25th IEEE International Conference on Image Processing (ICIP). pp. 1658–1662. IEEE (2018)