DAHYE KIM

School of Electrical and Electronic Engineering, Yonsei University, Seoul, Republic of Korea

Phone : (82-10) 5384-0901 ⋄ E-mail : dadaday@yonsei.ac.kr ⋄ https://dahye-kim.netlify.app/

RESEARCH INTERESTS

Computer Vision, Machine Learning, Video Understanding, Multi-modal Learning, Anomaly Detection, Out-of-distribution Detection, Learning with Limited Labeled Data

EDUCATION

Yonsei University, Seoul, Republic of Korea　　　　　　　　　　　*Sep. 2022 - Aug. 2024 (Expected)*
M.S. Candidate of Electrical & Electronic Engineering

- Advisor : Prof. Kwanghoon Sohn
- Cumulative GPA: 3.9/4.0

Yonsei University, Seoul, Republic of Korea　　　　　　　　　　　　　　　*Mar. 2020 - Aug. 2022*
B.S. of Electrical & Electronic Engineering

- Cumulative GPA: 4.0/4.0

Hongik University, Seoul, Republic of Korea　　　　　　　　　　　　　　　*Mar. 2018 - Feb. 2020*
Electronic & Electrical Engineering

- Cumulative GPA: 3.77/4.0

WORKING PAPERS

1. **Dahye Kim**, Jungin Park, Seongheon Park, and Kwanghoon Sohn (2024), "Towards Realistic Video Grounding via Temporal Token Merging", *IEEE/CVF Computer Vision and Pattern Recognition Conference (CVPR)*

2. Seongheon Park, Hyuk Kwon, **Dahye Kim**, Kwanghoon Sohn and Kibok Lee (2024), "Towards Imbalanced Open-World Semi-Supervised Learning", *IEEE/CVF Computer Vision and Pattern Recognition Conference (CVPR)*

PUBLICATIONS

1. **Dahye Kim**, Jungin Park, Jiyoung Lee, Seongheon Park, and Kwanghoon Sohn (2023), "Language-free Training for Zero-shot Video Grounding", *IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*

2. Seongheon Park, Hanjae Kim, Minsu Kim, **Dahye Kim**, and Kwanghoon Sohn (2023), "Normality Guided Multiple Instance Learning for Weakly Supervised Video Anomaly Detection", *IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*

WORK AND RESEARCH EXPERIENCES

Research Assistant, Digital Image Media Lab., Yonsei University　　　　　　　*Sep. 2022 - Present*

- Multi-modal Learning with Visual and Textual Representation
- Learning with Limited Labeled Data (Weakly-Supervised, Semi-Supervised, Zero-Shot Learning)
- Anomaly Detection and Open-Set Recognition for Computer Vision

Undergraduate Research Assistant, Digital Image Media Lab., Yonsei University *Jan. 2022 - Aug. 2022*

- Human-object Interaction (HOI) Detection
- Vision-Language Understanding
- Video Understanding

CONFERENCE PRESENTATIONS

1. **Dahye Kim**, Jungin Park, Jiyoung Lee, Seongheon Park, Kwanghoon Sohn, "Language-free Training for Zero-shot Video Grounding", *IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), Waikoloa, Hawaii, Jan. 2023, Oral Presentation.*

PROJECTS

1. **Korea Institute of Science and Technology (KIST)**　　　　　　　*Sep. 2022 - Aug. 2023*
   - Project Title: Deep Identification and Tracking of Missing Person in Heterogeneous CCTV
   - Covers: Object Detection, Object Tracking, Video Action Classification, Person Re-Identification, Anomaly Detection

2. **Project with Ministry of Science and ICT, Mid-Level Research**　　　　*Sep. 2022 - Aug. 2023*
   - Project Title: Development of Complex Situational Awareness and Prediction Technology through Multi-Modal Data Fusion and Social Artificial Intelligence
   - Covers: Multi-modal Learning, Video Understanding

3. **Project with Yonsei University-Yonsei Signature Research Cluster**　　　*Sep. 2022 - Aug. 2023*
   - Project Title: Development of Multimodal-based General-purpose Social Artificial Intelligence Technology
   - Covers: Multi-Modal Learning, Zero-Shot Learning, Meta-Learning

HONORS AND SCHOLARSHIPS

Honors

- High Honors Student, Yonsei University　　　　　　　　　　　*Fall, 2020 - Fall, 2021*
- Honors Student, Yonsei University　　　　　　　　　　　　　　　*Spring, 2020*

Scholarships

- Brain Korea 21 (BK21) Scholarship, National Research Foundation of Korea　*Sep. 2022 - Present*
- Automotive System IC Scholarship,　　　　　　　　　　　　　*Sep. 2022 - Present*
  Automotive System IC Fusion Human Resource Research Center
- Teaching Assistant Scholarship, Yonsei University　　　　　　　　　　*Fall, 2023*
  　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　Fall, 2022
- Graduate Student Research Assistant Scholarship, Yonsei University　　　　*Fall, 2022*
- Academic Excellence Scholarship, Yonsei University　　　　　　　　　　*Fall, 2021*
  　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　Spring, 2021
- National Science and Technology Scholarship, Korea Student Aid Foundation　*Fall, 2020*
- Academic Excellence Scholarship, Hongik University　　　　　　　*Fall, 2018 - Fall, 2019*

PATENTS

- Korean Patent No. 10-2023-0054355

TEACHING EXPERIENCES

- Signals and Systems　　　　　　　　　　　　　　　　　　　　　　*Fall, 2022*
- Introductory Digital Labs　　　　　　　　　　　　　　　　　　　　*Fall, 2023*

SELECTED COURSEWORK

Artificial Intelligence : Introduction Artificial Intelligence, Intelligent Control, Medical Imaging System and its Application for Artificial Intelligence, Special Topics for Deep Learning, Neural Network, Multimodal Deep Learning, Topics in Computer Vision.

Mathematics : Linear Algebra, Mathematical Statistics, Probability and Random Variables, Random Process, Optimization Theory, Information Theory.

Programming : Data Structure and Algorithms, Digital Image Processing, Operating Systems.

SKILL SET

Language and Tools : Python, MATLAB, C++, LaTeX.

Libraries : Pytorch, TensorFlow, Keras, Scikit-Learn, Numpy, Jupyter, etc.

APPENDIX: SUMMARY OF RESEARCH

PUBLICATIONS

Language-free Training for Zero-shot Video Grounding

Dahye Kim , Jungin Park, Jiyoung Lee, Seongheon Park, Kwanghoon Sohn

Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. 2023.

Abstract : Given an untrimmed video and a language query depicting a specific temporal moment in the video, video grounding aims to localize the time interval by understanding the text and video simultaneously. One of the most challenging issues is an extremely time- and cost-consuming annotation collection, including video captions in a natural language form and their corresponding temporal regions. In this paper, we present a simple yet novel training framework for video grounding in the zero-shot setting, which learns a network with only video data without any annotation. Inspired by the recent language-free paradigm, i.e. training without language data, we train the network without compelling the generation of fake (pseudo) text queries into a natural language form. Specifically, we propose a method for learning a video grounding model by selecting a temporal interval as a hypothetical correct answer and considering the visual feature selected by our method in the interval as a language feature, with the help of the well-aligned visual-language space of CLIP. Extensive experiments demonstrate the prominence of our language-free training framework, outperforming the existing zero-shot video grounding method and even several weakly-supervised approaches with large margins on two standard datasets.

Normality Guided Multiple Instance Learning for Weakly Supervised Video Anomaly Detection

Seongheon Park, Hanjae Kim, Minsu Kim, **Dahye Kim**, Kwanghoon Sohn

Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. 2023.

Abstract : Weakly supervised Video Anomaly Detection (wVAD) aims to distinguish anomalies from normal events based on video-level supervision. Most existing works utilize Multiple Instance Learning (MIL) with ranking loss to tackle this task. These methods, however, rely on noisy predictions from a MIL-based classifier for target instance selection in ranking loss, degrading model performance. To overcome this problem, we propose Normality Guided Multiple Instance Learning (NG-MIL) framework, which encodes diverse normal patterns from noise-free normal videos into prototypes for constructing a similarity-based classifier. By ensembling predictions of two classifiers, our method could refine the anomaly scores, reducing training instability from weak labels. Moreover, we introduce normality clustering and normality guided triplet loss constraining inner bag instances to boost the effect of NG-MIL and increase the discriminability of classifiers. Extensive experiments on three public datasets (ShanghaiTech, UCF-Crime, XD-Violence) demonstrate that our method is comparable to or better than existing weakly supervised methods, achieving state-of-the-art results.