Mengmeng Xu | Curriculum Vitae

KAUST H4201-205B, Thuwal, 23955, KSA

(+44)-7868-244-953 • xu.frost@gmail.com • frostinassiky

"Interdisciplinary Engineer with Entrepreneurial Mindset"

RESEARCH INTERESTS

I am focusing on problems that arise in image and video understanding. Particularly, I am interested in the representation learning of long untrimmed videos via self-supervised learning with novel model architectures, and video localization tasks such as temporal action/object localization and language video grounding.

SKILLS

- Python/Matlab/C/C++
- PyTorch/TensorFlow

WORK EXPERIENCE

Facebook UK Limited **London, UK**
Research Scientist Intern in Facebook AI *2022 Mar.- Jul.*

Amazon Development Center Germany GmbH **Berlin, Germany**
Applied Scientist Intern in Amazon CVNA group *2021 Aug.- 2022 Jan.*

Samsung Electronics (UK) Ltd **Cambridge, UK**
Research Intern in Samsung AI Center (SAIC) *2020 Sep.- 2021 Mar.*

EDUCATION

King Abdullah University of Science and Technology **Thuwal, KSA**
MS-PhD in Electrical and Computer Engineering, GPA: 3.88/4.00 *2017 - 2023*

Zhejiang University **China**
B.S. in Opto-Elctronics Information Science and Engineering, GPA 3.92/4.00 *2013 - 2017*

SELECTED PROJECTS

Egocentric Visual Query 2D Localization *2022 Mar. - Present*
We study the VQ2D problem in Ego4D, which aims to retrieve objects from the video in the first-person view. This task is challenging due to the limited annotation, high diversity in object classes, and abnormal view transformations.
- We reduced model's false positive rate by leveraging noisy, blurry, and unlabeled background frames.
- We managed to scale the training pipeline effectively and efficiently, and got the first prize in CVPR workshop.
- We designed a data augmentation method considering object position change in different egocentric views.

Egocentric 4D Perception (EGO4D) *2019 Dec. - 2021 Oct.*
We introduce Ego4D, a massive-scale egocentric video dataset and benchmark suite. It offers 3,025 hours of daily-life activity video captured by 855 unique camera wearers from 74 worldwide locations and 9 different countries.
- Collaborated with 88 researchers in an international consortium, we dramatically increases the scale of egocentric data publicly available, making it more than 20x greater than any other data set in terms of hours of footage.
- The dataset is diverse in its geographic coverage, scenarios, participants, and captured modalities.
- It will enable many applications in egocentric perception from human-robot interactions to virtual personal assistants.

Low-Fidelity End-to-End Video Encoder Pre-training for TAL *2021 Feb. - Jun.*
End-to-end video encoder pre-training for temporal action localization is not operable subject to the GPU memory constraints, due to the prohibitive computational cost in processing long untrimmed videos.
- We reduce the mini-batch composition so that end-to-end optimization for the video encoder becomes operable.
- LoFi-TAL favourably solves the task discrepancy problem and providing more effective feature representations.
- LoFi-TAL with lightweight ResNet18 in a RGB stream surpasses RGB+optical-flow two-stream ResNet50 models.

Boundary-sensitive Pre-training for Temporal Localization in Videos *2020 Sep. - 2021 Mar.*
Most existing models for temporal localization tasks are pre-trained on video classification tasks. The domain gap between action recognition and localization can be addressed by a temporal boundary datasets.

- For the first time, we investigate pre-training for localization by introducing a novel boundary-sensitive pretext task.
- We propose to synthesize temporal boundaries in existing video classification datasets to help localize action.
- Extensive experiments show that the proposed BSP is superior and complementary to the existing action classification based pre-training counterpart, and achieves new state-of-the-art performance on **several** temporal localization tasks.

Sub-Graph Localization for Temporal Action Detection *2019 Jan - Nov*
Recent studies show that context can be used as a clue to help understanding action before or after the snippet. Thus, we formulated action localization task to a sub-graph detection problem, solved by graph convolutional network.

- A multi-graph convolutional layer is designed to progressively represent the video snippet by its adaptive semantics.
- The method achieves SOTA performance on two large-scale video benchmarks for human activity localization.

SELECTED PUBLICATIONS/PREPRINTS

Ego4D: Around the World in 3,000 Hours of Egocentric Video **2022**
Kristen Grauman, ···, Mengmeng Xu, ···, Jitendra Malik,* (*= key contributor) *CVPR*

Low-Fidelity E2E Video Encoder Pre-training for Temporal Action Localization **2021**
Mengmeng Xu, ···, Xiatian Zhu, Bernard Ghanem, Brais Martinez *NeurIPS*

Boundary-sensitive Pre-training for Temporal Localization in Videos **2021**
Mengmeng Xu, ···, Xiatian Zhu, Li Zhang, Bernard Ghanem, Tao Xiang *ICCV*

Relation-aware Video Reading Comprehension forTemporal Language Grounding **2021**
Jialin Gao, Xin Sun, Mengmeng Xu, Xi Zhou, Bernard Ghanem *EMNLP*

G-TAD: Sub-Graph Localization for Temporal Action Detection **2020**
Mengmeng Xu, Chen Zhao, David S. Rojas, Ali Thabet, Bernard Ghanem *CVPR*

BAOD: Budget-Aware Object Detection **2021**
Alejandro Pardo, Mengmeng Xu*, ···, Bernard Ghanem* *CVPR Workshop,* ***Best Paper***

Missing Labels in Object Detection **2019**
Mengmeng Xu, Yancheng Bai, B Ghanem *CVPR Workshop*

Object Detection Using Multiple Level Annotations **2019**
Mengmeng Xu *Master Thesis, KAUST*

ACADEMIC/PROFESSIONAL ACTIVITIES

Academic Reviewer ***2019 - Present***
ICCV, AAAI, CVPR, ECCV, NeurIPS, TPAMI

Program Chair ***2019 - Present***
ActivityNet Challenge workshop in CVPR19, CVPR20, CVPR21, CVPR22

Teaching Assistant ***2019 - 2020***
CS390DD: Special Topics in Machine Learning; AMCS211: Numerical Optimization *KAUST*

Certificates and Nano-degrees ***2016 - 2020***
NLP Nanodegree, Deep Learning Specialization, Machine Learning. *Coursera and Udacity*

AWARDS/HONORS

- CEMSE Dean's List Award, *KAUST* 2022
- **First Prize** in the Visual Query 2D Localization task in Ego4D Challenge, *CVPR Workshop* 2022
- **First Prize** in the Innovation Challenge of LOng-form VidEo Understanding, *CVPR Workshop* 2021
- **Outstanding Reviewer**, CVPR 21 2021
- **Outstanding Graduates** of Zhejiang University 2017
- **Chu Kochen Honors Program** for Advanced Engineering Education, *ZJU.* 2014-2017
- **Gold Medal**, International Genetically Engineered Machine Competition (iGEM), *Boston, USA* 2016
- **National Scholarship**, *Zhejiang University* 2014
- **First prize** (×2) in Olympic Competition of Physics, Olympic Competition of Mathematics, *China* 2012