Kumara Kahatapitiya

Department of Computer Science
Stony Brook University
Stony Brook, NY 11794

cs.stonybrook.edu/~kkahatapitiy
kkahatapitiy@cs.stonybrook.edu

Research Interests

Video Understanding, Video Generation/Editing, Diffusion Models, Multi-modal Models

Education

Stony Brook University, Stony Brook, NY
Ph.D, Computer Science, 2019 - 2024 (expected)
GPA: 4.0/4.0, Thesis: Video Representations: A View on Inference and Label Efficiency

University of Moratuwa, Sri Lanka
B.Sc. (Hons.), Electronic & Telecommunication Engineering, 2014 - 2018
GPA: 4.05/4.20

Employment

Stony Brook University, Stony Brook, NY
Research Assistant, Summer 2020 - Present
Teaching Assistant, Fall 2019 - Spring 2020

Meta Platforms, Inc., Menlo Park, CA
Research Scientist Intern, Summer 2024

Qualcomm AI Research, Amsterdam, Netherlands
Research Intern, Summer 2023

Google DeepMind, New York City, NY
Student Researcher, Spring 2022 - Spring 2023

Wormpex AI Research, Bellevue, WA
Research Intern, Summer 2021

Publications

Preprints

1. Adaptive Caching for Faster Video Generation with Diffusion Transformers
Kumara Kahatapitiya, Haozhe Liu, Sen He, Ding Liu, Menglin Jia, Chenyang Zhang, Michael S. Ryoo, Tian Xie
arXiv 2024 [project-page] [preprint] [code]

2. MarDini: Masked Auto-Regressive Diffusion for Video Generation at Scale
Haozhe Liu, Shikun Liu, Zijian Zhou, Mengmeng Xu, Yanping Xie, Xiao Han, Juan C. Pérez, Ding Liu, Kumara Kahatapitiya, Menglin Jia, Jui-Chieh Wu, Sen He, Tao Xiang, Jürgen Schmidhuber, Juan-Manuel Pérez-Rúa
arXiv 2024 `[project-page] [preprint]`

3. Understanding Long Videos in One Multimodal Language Model Pass
Kanchana Ranasinghe, Xiang Li, Kumara Kahatapitiya, Michael S. Ryoo
arXiv 2024 `[project-page] [preprint] [code] [webinar]`

Peer-reviewed Articles

4. Language Repository for Long Video Understanding
Kumara Kahatapitiya, Kanchana Ranasinghe, Jongwoo Park, Michael S. Ryoo
NeurIPS 2024 workshops `[paper] [code] [webinar]`

5. Too many frames, not all useful: Efficient Strategies for Long-form Video QA
Jongwoo Park, Kanchana Ranasinghe, Kumara Kahatapitiya, Wonjeong Ryoo, Donghyun Kim, Michael S. Ryoo
NeurIPS 2024 workshops `[paper] [code] [webinar]`

6. LLaRA: Supercharging Robot Learning Data for Vision-Language Policy
Xiang Li, Cristina Mata, Jongwoo Park, Kumara Kahatapitiya, Yoo Sung Jang, Jinghuan Shang, Kanchana Ranasinghe, Ryan Burgert, Mu Cai, Yong Jae Lee, Michael S. Ryoo
CoRL 2024 workshops `[paper] [code]`

7. Object-Centric Diffusion for Efficient Video Editing
Kumara Kahatapitiya, Adil Karjauv, Davide Abati, Yuki M. Asano, Fatih Porikli, Amirhossein Habibian
ECCV 2024 `[project page] [paper]`

8. VicTR: Video-conditioned Text Representations for Activity Recognition
Kumara Kahatapitiya, Anurag Arnab, Arsha Nagrani, Michael S. Ryoo
CVPR 2024 `[paper] [poster] [talk]`

9. Grafting Vision Transformers
Jongwoo Park, Kumara Kahatapitiya, Donghyun Kim, Shivchander Sudalairaj, Quanfu Fan, Michael S. Ryoo
WACV 2024 `[paper] [poster]`

10. SWAT: Spatial Structure Within and Among Tokens
Kumara Kahatapitiya, Michael S. Ryoo
IJCAI 2023 `[paper] [code] [slides]`

11. Token Turing Machines
Michael S. Ryoo, Keerthana Gopalakrishnan, Kumara Kahatapitiya, Ted Xiao, Kanishka Rao, Austin Stone, Yao Lu, Julian Ibarz, Anurag Arnab
CVPR 2023 `[paper] [code] [teaser]`

12. Weakly-guided Self-supervised Pretraining for Temporal Activity Detection
Kumara Kahatapitiya, Zhou Ren, Haoxiang Li, Zhenyu Wu, Michael S. Ryoo, Gang Hua
AAAI 2023 `[paper] [code] [talk] [poster]`

13. StARformer: Transformer with State-Action-Reward Representations for Visual Reinforcement Learning
Jinghuan Shang, Kumara Kahatapitiya, Xiang Li, Michael S. Ryoo
ECCV 2022 `[paper] [code] [poster]`
T-PAMI `[paper]`

14. MS-TCT: Multi-Scale Temporal ConvTransformer for Action Detection
Rui Dai, Srijan Das, Kumara Kahatapitiya, Michael S. Ryoo, Francois Bremond
CVPR 2022 `[paper] [code]`

15. Swift: Adaptive Video Streaming with Layered Neural Codecs
Mallesham Dasari, Kumara Kahatapitiya, Samir Das, Aruna Balasubramanian, Dimitris Samaras
NSDI 2022 [paper] [code] [slides]

16. Coarse-Fine Networks for Temporal Activity Detection in Videos
Kumara Kahatapitiya, Michael S. Ryoo
CVPR 2021 [paper] [code] [talk] [poster]

17. Exploiting the Redundancy in Convolutional Filters for Parameter Reduction
Kumara Kahatapitiya, Ranga Rodrigo
WACV 2021 [paper] [code] [talk]

18. Feature-dependent Cross-Connections in Multi-Path Neural Networks
Dumindu Tissera, Kasun Vithanage, Rukshan Wijesinghe, Kumara Kahatapitiya, Subha Fernando, Ranga Rodrigo
ICPR 2020 [paper]

19. Context-Aware Automatic Occlusion Removal
Kumara Kahatapitiya, Dumindu Tissera, Ranga Rodrigo
ICIP 2019 [paper] [code]

Distinctions

Finalist (1/30) for Adobe Research Fellowship, 2022

Nomination for Microsoft Research PhD Fellowship and Google PhD Fellowship, 2022

Gold Medalist, Electronic & Telecommunication Engineering, University of Moratuwa, 2018

Dean's List Recognition (7/8 semesters), University of Moratuwa, 2014-2018

National Inter-University Statistics Champions, Sri Lanka, 2015

Presidential Recognition for the excellence in national examinations, Sri Lanka, 2009, 2012

Academic Service

Reviewer at CVPR, ICCV, ECCV, T-PAMI, IJCV, NeurIPS, ICLR, ICML, AAAI, WACV

Open-source Projects

X3D-Multigrid [code] , Optimal Transport in NumPy [code]

References

Michael S. Ryoo, SUNY Empire Innovation Associate Professor, Stony Brook University
mryoo@cs.stonybrook.edu

Dimitris Samaras, SUNY Empire Innovation Professor, Stony Brook University
samaras@cs.stonybrook.edu

Amirhossein Habibian, Director of Engineering, Qualcomm AI Research
ahabibian@qti.qualcomm.com

Dmitry Kalashnikov, Staff Software Engineer, Google DeepMind
dkalashnikov@google.com