

# KANCHANA RANASINGHE

kranasinghe@cs.stonybrook.edu | <http://kahnchana.github.io>

## EDUCATION

---

**Stony Brook University, NY, USA**

Aug 2021 - Present

*PhD in Computer Science*

**University of Moratuwa, Sri Lanka**

Dec 2015 - Jan 2020

*BSc in Engineering; Awarded Most Outstanding Graduatand of the Year*

Senior Thesis Title: “Realtime Multi-Object Tracking and Pixelwise Segmentation”

## RESEARCH EXPERIENCE

---

**Salesforce Research, Palo Alto, USA - Research Intern**

June 2025 - Present

- Learning Language-Motion Relations from Human Activity Videos
- Unified Generative Models for Robot Control

**Google Research, NYC, USA - Student Researcher**

March 2024 - Dec 2024

- Continuous-space CRF formulation for modeling natural images
- Integrating CRF with latent diffusion for efficient image generation

**Meta AI Research, NYC, USA - Research Scientist Intern**

May 2023 - Dec 2023

- Spatial reasoning in multi-modal large language models (CVPR '24)
- Localization awareness in video-language models

**Apple MLR, Cupertino, USA - Machine Learning Research Intern**

May 2022 - Sep 2022

- Multi-modal self-supervised representation learning (ICCV '23)
- Interpretability and robustness of vision language models

**MBZUAI, Abu Dhabi, UAE - Research Assistant**

Nov 2020 - Aug 2021

- Representation learning: contrastive losses, self-supervised objectives (ICCV '21)
- Interpretability, robustness, and adversarial attacks for vision transformers (NeurIPS '21, ICLR '22)
- Generative modelling for multi-modal output spaces (ICLR '21)

**Wenn ASA, Stavanger, Norway - Machine Learning Engineer (remote)**

Jan 2019 - Oct 2020

- Leading team of three associate data scientists for vehicle damage detection project
- Focus on efficient mobile implementations, instance segmentation models, and active learning

**FiveAI, Cambridge, UK - Research Intern**

June 2018 - Dec 2018

- Perception team of self-driving startup
- 3D orientation estimation: improve occluded object handling in videos with synthetic data

## SELECTED AWARDS

---

**Outstanding Demo Award** at CVPR 2023

Burgert, Ranasinghe, Li, Ryoo, “*Diffusion Illusions: Hiding Images in Plain Sight*”

**Most Outstanding Graduatand of the Year**

Class of 2020 at University of Moratuwa, Sri Lanka

## PUBLICATIONS - CONFERENCE

---

**Understanding Long Videos with Multimodal Language Models**

ICLR, 2025

K Ranasinghe, X Li, K Kahatapitiya, M S Ryoo

**LLaRA: Supercharging Robot Learning Data for Vision-Language Policy** ICLR, 2025  
X Li, C Mata, J Park, K Kahatapitiya, Y Jang, J Shang, K Ranasinghe, R Burgert, M Cai, Y Lee, M S Ryoo

**Language Repository for Long Video Understanding** ACL Findings, 2025  
K Kahatapitiya, K Ranasinghe, J Park, M S Ryoo

**Unsupervised Domain Adaptive Segmentation using Domain-Agnostic Text** ECCV, 2024  
C Mata, K Ranasinghe, M S Ryoo

**Learning to Localize Objects Improves Spatial Reasoning in Visual-LLMs** CVPR, 2024  
K Ranasinghe, S N Shukla, O Poursaeed, M S Ryoo, T Y Lin

**Too Many Frames, not all Useful: Efficient Strategies for Long-Form Video QA** NeurIPS-W, 2024  
J Park, K Ranasinghe, K Kahatapitiya, W Ryoo, D Kim, M S Ryoo

**Diffusion Illusions: Hiding Images in Plain Sight** SIGGRAPH, 2024  
R Burgert, X Li, A Leite, K Ranasinghe, M Ryoo

**Hierarchical Text-to-Vision Self Supervised Alignment for Improved Histopathology Representation Learning** MICCAI 2024  
H Watawana, K Ranasinghe, T Mahmood, M Naseer, S Khan, F S Khan

**Language-based Action Concept Spaces Improve Video SSL** NeurIPS, 2023  
K Ranasinghe, M S Ryoo

**Perceptual Grouping in Contrastive Vision-Language Models** ICCV, 2023  
K Ranasinghe, B McKinzie, S Ravi, Y Yang, A Toshev, J Shlens

**Peekaboo: Text to Image Diffusion Models are Zero-Shot Segmentors** CVPR workshop, 2023  
R Burgert, K Ranasinghe, X Li, M S Ryoo

**Self-supervised Video Transformers** CVPR, 2022 (oral)  
K Ranasinghe, M Naseer, S Khan, F Khan, M S Ryoo

**On Improving Adversarial Transferability of Vision Transformers** ICLR, 2022 (spotlight)  
M Naseer, K Ranasinghe, S Khan, F Khan, F Porikli

**Intriguing Properties of Vision Transformers** NeurIPS, 2021 (spotlight)  
M Naseer, K Ranasinghe, S Khan, M Hayat, F Khan, M Yang  
(*listed among top 15 most cited papers at NeurIPS 2021 - link*)

**Orthogonal Projection Loss** ICCV, 2021  
K Ranasinghe, M Naseer, M Hayat, S Khan, F Khan

**Conditional Generative Modeling via Learning the Latent Space** ICLR, 2021  
S Ramasinghe, K Ranasinghe, S Khan, N Barnes, S Gould

**Bipartite Conditional Random Fields for Panoptic Segmentation** BMVC, 2020 (oral)  
S. Jayasumana, K Ranasinghe, M. Jayawardhana, S. Liyanaarachchi and H. Ranasinghe

**Micro Actions and Deep Static Features for Activity Recognition** DICTA, 2017  
S. Ramasinghe, J. Rajasegaran, V. Jayasundara, K. Ranasinghe, R. Rodrigo and A. Pasqual

---

#### PUBLICATIONS - PREPRINTS

**Pixel Motion as Universal Representation for Robot Control**  
K Ranasinghe, X Li, C Mata, J Park, M S Ryoo

**Blip-3-Video: You Only Need 32 Tokens to Represent a Video Even in VLMs**  
Michael S Ryoo, Honglu Zhou, Shrikant Kendre, Can Qin, Le Xue, Manli Shu, Jongwoo Park, Kanchana Ranasinghe, Silvio Savarese, Ran Xu, Caiming Xiong, Juan Carlos Niebles

**Test-Time Optimization for Domain Adaptive Open Vocabulary Segmentation**  
U Silva, D Samaraweera, S Wanigathunga, K Kariyawasam, K Ranasinghe, M Naseer, R Rodrigo

**LatentCRF: Continuous CRF for Efficient Latent Diffusion**  
K Ranasinghe, S Jayasumana, A Veit, A Chakrabarti, D Glasner, M S Ryoo, S Ramalingam, S Kumar

## PUBLICATIONS - JOURNAL

---

**Combined Static & Motion Features for Deep-Networks Based Activity Recognition in Videos**, S. Ramasinghe, J. Rajasegaran, V. Jayasundara, K Ranasinghe, R. Rodrigo and A. A. Pasqual, IEEE Transactions on Circuits and Systems for Video Technology, vol. 29, Sept. 2019.

## PATENTS

---

Kanchana Ranasinghe, Muhammad Muzammal Naseer, Salman Khan, Fahad Khan, **“System and Method for Self-supervised Video Transformer.”** US Patent US20240169692A1 filed on 11/21/2022.

## INVITED TALKS

---

*“Visual Spatial Reasoning with Multimodal LLMs”*, presented (remotely) at Deep Learning Session in Radiology Department, Penn State University, PA (Jan 2025).

*“MVU - an LLM-based framework for solving Long Video Question Answering benchmarks”*, presented (remotely) at Multimodal Weekly organized by Twelve Labs (Aug 2024).

*“Learning to Localize Objects Improves Spatial Reasoning in Visual-LLMs”*, presented at NYC Vision RG, Google Research, NYC (Apr 2024).

*“Perceptual Grouping in Contrastive Vision-Language Models”*, presented at Electronics and Telecommunications Department, University of Moratuwa, Sri Lanka (Jan 2024).

*“Perceptual Grouping in Contrastive Vision-Language Models”*, presented (remotely) at Computer Vision Talks series, ETS Montreal (Oct 2023).

## TEACHING

---

### **Stony Brook University, NY, USA**

- Teaching assistant for CSE 101 course on Introduction to Digital Intelligence
- Teaching assistant for CSE 215 course on Foundations of Computer Science course
- Supporting for graduate level computer vision seminar CSE 656

### **University of Moratuwa, Sri Lanka**

- Teaching assistant for undergraduate course on Introduction to Computer Vision
- Supporting for undergraduate level computer vision seminar

## SERVICE

---

### **Reviewing**

- Conference reviewer: CVPR, ICCV, ECCV, NeurIPS, ICLR, ICML, ICRA 2022 - present
- Journal reviewer: IJCV, TCSVT 2021 - present

### **Mentoring**

- Mentor for undergraduate computer science students at Women in Engineering in Sri Lanka 2023
- Mentor for machine learning competition teams at University of Moratuwa, Sri Lanka 2023
- Mentor undergraduates from University of Moratuwa, Sri Lanka in computer vision research projects and academic writing 2023 - present

## OTHER AWARDS

---

**National Merit Scholarship** - Ranked 13th (2.83 z-score) in Sri Lanka at GCE A/L Examination 2014

**World Rank 296 / National Top 6** - International Mathematical Olympiad (IMO), Colombia 2013