

PRITAM SARKAR

Toronto, Canada | [google scholar](#) | [github](#) | [linkedin](#) | pritam.sarkar@queensu.ca | www.pritamsarkar.com

Final-year PhD candidate with prior industry experience (internships at Google and Borealis AI, along with 3 years as a Software Engineer), solid research background (NeurIPS/ICLR/AAAI/ICASSP), and a perfect academic record (GPA 4.3/4.3), looking for a researcher role.

RESEARCH INTERESTS

My current research focus is **video and multimodal learning**.

Specifically, it encompasses: multimodal learning with vision, language, and audio; video understanding; representation learning; self-supervised and unsupervised learning; multimodal LLMs; large vision-language models; foundation models; and computer vision.

EXPERIENCE

Queen's University

Research Assistant (PhD)

Supervisor: [Ali Etemad](#)

2020 - Present

Kingston, Canada

During this time I am also affiliated with the **Vector Institute** (2021 - Present).

- Currently, I am working on self-alignments of multimodal LLMs for video understanding.
- For the first time studied the behaviour of popular video self-supervised methods in response to various forms of natural distribution shift, uncovering a series of intriguing findings and interesting behaviors.
- Introduced XKD to improve alignment between audio and visual modalities in video representation learning.
- Introduced CrissCross to learn generalized representations leveraging the asynchronous relationships between audio and visual modalities.
- Introduced AVCAffe, the largest recorded affective video dataset for human behavioural understanding.

Google

Student Researcher/Research Intern

Hosts: [Sayna Ebrahimi](#) and [Sercan Ö. Arık](#).

Introduced phrase-level alignment, to mitigate vision-language hallucinations in multimodal LLMs while preserving their ability on general tasks, unlike existing finetuning methods.

Fall 2023

Sunnyvale, USA

Borealis AI

Machine Learning Research Intern

Host: [Fredrick Tung](#)

Addressed the limitation of standard augmentation techniques in event sequence time-series data by developing AugESeq, a conditional diffusion model that generates realistic augmented event sequences for improved representation learning.

Fall 2022

Toronto, Canada

Queen's University

Research Assistant (MASC)

Supervisor: [Ali Etemad](#)

2018 - 2020

Kingston, Canada

- I was the first to introduce self-supervised learning with ECG (Electrocardiograms).
- I was the first to introduce PPG (Photoplethysmogram) to ECG translation using deep learning for continuous cardiac activity monitoring.

Queen's University

Guest Lecturer/Lead Teaching Assistant/Teaching Assistant

I worked for the following courses: Artificial Intelligence & Interactive Systems (ELEC 872 in F'23), Artificial Intelligence (ELEC 472 in W'25, W'23, W'22, W'21, S'19), Electronics I (ELEC-252 in F'20), Introduction to Computer Programming for Engineers (APSC-143 in F'19)

2018 - Present

Kingston, Canada

Infosys Ltd.

Senior System Engineer

I worked on Oracle Cloud and SQL.

2017 - 2018

Bangalore, India

Tech Mahindra Ltd.

Software Engineer

I worked on Oracle Cloud and SQL.

2015 - 2017

Hyderabad, India

EDUCATION

Doctor of Philosophy (PhD)

Dept. of Electrical and Computer Engineering, Queen's University

Thesis topic: Video and Multimodal Learning

Advisor: [Ali Etemad](#)

GPA: 4.3/4.3

May 2020 - Apr 2025 (expected)

Kingston, Canada

Master Applied Science (MASc)

Dept. of Electrical and Computer Engineering, Queen's University

Thesis title: Self-Supervised ECG Representation Learning for Affective Computing. [\[Link to thesis\]](#)

Advisor: [Ali Etemad](#)

GPA: 3.8/4.3

Sept 2018 - Apr 2020

Kingston, Canada

Bachelor of Technology (B.Tech)

Dept. of Electrical Engineering, West Bengal University of Technology

Rank 4 of 150 in graduating class of Electrical Engineering.

GPA: 8.84/10

Aug 2011 - Jul 2015

Kolkata, India

SELECTED PUBLICATIONS

As per [google scholar](#), my publications have total citations of approx. 725. Please find the full list [here](#).

10. P. Sarkar, S. Ebrahimi, A. Etemad, A. Beirami, S. Arik, T. Pfister, "Data-Augmented Phrase-Level Alignment for Mitigating Object Hallucination", *ICLR* 2025. [\[Paper\]](#)
9. P. Sarkar, A. Etemad, "XKD: Cross-modal Knowledge Distillation with Domain Alignment for Video Representation Learning", *AAAI* 2024. [\[Paper\]](#)
8. P. Sarkar, A. Beirami, A. Etemad, "Uncovering the Hidden Dynamics of Video Self-supervised Learning under Distribution Shifts", *NeurIPS* 2023. [Spotlight \[Paper\]](#)
7. P. Sarkar, A. Etemad, "Self-supervised Audio-Visual Representation Learning with Relaxed Cross-Modal Synchronicity", *AAAI*, 2023. [Oral \[Paper\]](#)
6. P. Sarkar, A. Posen, A. Etemad, "AVCAffe: A Large Scale Audio-Visual Dataset of Cognitive Load and Affect for Remote Work", *AAAI*, 2023. [\[Paper\]](#)
5. D. Shome, P. Sarkar, A. Etemad, "Region-Disentangled Diffusion Model for High-Fidelity PPG-to-ECG Translation", *AAAI* 2024. [\[Paper\]](#)
4. P. Sarkar, A. Etemad, "Method and Apparatus for Generating an Electrocardiogram from a Photoplethysmogram", *US20230363655A1*. [Patent \[Link\]](#)
3. P. Sarkar, A. Etemad, "CardioGAN: Attentive Generative Adversarial Network with Dual Discriminators for Synthesis of ECG from PPG", *AAAI*, 2021. [\[Paper\]](#)
2. P. Sarkar, A. Etemad, "Self-supervised ECG Representation Learning for Emotion Recognition", *IEEE Transactions on Affective Computing*, 2020. [\[Paper\]](#) 300+ citations
1. P. Sarkar, A. Etemad, "Self-supervised Learning for ECG-based Emotion Recognition", *ICASSP*, 2020. [Oral \[Paper\]](#) 100+ citations

TECHNICAL SKILLS

Deep Learning Frameworks: [PyTorch](#), TensorFlow, Keras

Programming Languages: [Python](#), MATLAB, C, SQL

ACADEMIC ACHIEVEMENTS/AWARDS

- [First prize](#) in IEEE Research Excellence Award (PhD), 2023, IEEE Kingston Section.
- [Best Poster](#) Award at Robotics and AI Symposium at Ingenuity Labs, 2023. Title: *Cardiac Insights On-the-Go: Inexpensive Continuous ECG Monitoring from PPG Using Diffusion Models*
- [Honourable Mention](#) at FEAS Research Symposium at Queen's University, 2022. Title: *The First Large-Scale Audio-Visual Dataset of Cognitive Load and Affect for Remote Work*
- [Honourable Mention](#) at Robotics and AI Symposium at Ingenuity Labs, 2022. Title: *The First Large-Scale Audio-Visual Dataset of Cognitive Load and Affect for Remote Work*
- [Best Poster](#) Award at Robotics and AI Symposium at Ingenuity Labs, 2021. Title: *Toward Wearables of the Future: Affordable Acquisition of Continuous ECG with Deep Learning*
- Postgraduate Affiliate Award, Vector Institute, 2021 - 2023.

- Graduate Research Fellowship, Queen's University, 2020 - 2025.
- Graduate Research Scholarship, Queen's University, 2019 - 2020.

INVITED TALKS

- June 2024 at Google: Data-Augmented Phrase-Level Alignment for Mitigating Object Hallucination
- July 2023 at Ingenuity Labs: Learning without Human Supervision
- January 2023 at Borealis AI: Augmentation Improves Event Sequence Prediction

ACADEMIC SERVICE

Mentorship

- Seth Grief-Albert, Bachelor at ECE, Queen's University, Summer 2024.
- Vishal Narnaware, Visiting Student at University of Cambridge, co-mentored with Nikhil Churamani, 2023-2024.
- Debaditya Shome, MASc at ECE, Queen's University, 2022-2023.
- Aaron Posen, Bachelor at ECE, Queen's University, 2021-2022.
- Rachel Phinnemore, Bachelor at CS, Queen's University, 2020-2021.

Reviewing/PC Member

I regularly review for the following venues:

- NeurIPS, ICLR, AAAI, CVPR, ICCV, ECCV, ICML, ICASSP, ACII
- IEEE Transactions on - PAMI, Affective Computing, Artificial Intelligence

Organizing workshops and conferences

- Session chair for computer vision tracks at AAAI 2023.
- **Co-organizer** of AAAI 2023 Workshop on Representation Learning for Responsible Human-centric AI ([R2HCAI](#)).
- **Co-organizer** of AAAI 2022 Workshop on Human-centric Self-supervised Learning ([HCSSL](#)).
- Volunteer at AI/GI/CRV Conference, 2019.

Others

- Student Rep. in Graduate Studies Academic Advisory Committee, Dept. of ECE, Queen's University, 2020 - 2021.
- PhD Rep. at Graduate Electrical and Computer Engineering student council, Queen's University, 2020 - 2021.

REFERENCES

- Dr. Ali Etemad, Associate Professor at Queen's University, email: ali.etemad@queensu.ca
- Dr. Ahmad Beirami, Research Scientist at Google DeepMind, email: beirami@google.com

CREATIVE INTERESTS

Photography and short-film-making.