

# Jaejun Brandon Lee

COMPUTER VISION · NATURAL LANGUAGE PROCESSING · SPEECH RECOGNITION · ROBOTICS  
☎ (+1) 204-698-1771 | ✉ ljj7975@gmail.com | 📱 ljj7975 | 📺 ljj7975 | 🏠 ljj7975.github.io

## Introduction

Drawing from seven years of research experience across computer vision, speech recognition, natural language processing, and robotics, I believe that multi-modal representation learning is not only effective in capturing knowledge about our world but also in controlling and interpreting a model's internal representations. With that in mind, **I am particularly interested in building a vision system that perceives the world as humans do, leveraging the effective alignment of visual and textual data.**

## Education

### University of Waterloo

MASTER OF MATHEMATICS IN COMPUTER SCIENCE, ADVISOR: DR. JIMMY LIN

Waterloo, ON, Canada

2020

- Thesis title: In-Browser Personalization for Ubiquitous Keyword Spotting
- Research area: Information retrieval and deep learning (natural language processing & speech recognition)

### University of Waterloo

BACHELOR OF COMPUTER SCIENCE

Toronto, ON, Canada

2018

- Completed co-operative program and graduated with distinction
- Recipient of President's Scholarship and Faculty of Mathematics Scholarship

## Scholarships

2013-2018	<b>Scholarship</b> , Faculty of Mathematics Scholarship, University of Waterloo	value of CAD 25,000
2013	<b>Scholarship</b> , President's Scholarship, University of Waterloo	value of CAD 2,000
2013	<b>Scholarship</b> , Winnipeg North Rotary Club Scholarship	value of CAD 5,000
2013	<b>Scholarship</b> , Loblaw Scholarship	value of CAD 1,500

## Experience

### Research Scientist

EPSON CANADA - COMPUTER VISION & ROBOTICS LAB

Markham, ON, Canada

Mar. 2023 - Present

- Conducting research to develop an interpretable 2D object and attribute detection system tailored for non-expert users.
  - Designed a two-stage adapter-based domain adaptation method that utilizes a contrastive loss focused on hard-negative samples.
  - Enabling effective verification of the model's representations by establishing a strong understanding of context-agnostic foundational concepts.
  - Developing a plug-and-play, non-parametric detection head for vision-language models.
- Proposed and developed a prompt-based 2D keypoint detection pipeline that eliminates keypoint-specific training needs while remaining robust in multi-instance scenarios (a first-author paper under review for TAI [14]).

### Lead Research Scientist

ROBOEYE.AI

Toronto, ON, Canada

Jul. 2021 - Mar. 2023

- Developed a real-time (<1 sec) 6D pose estimation pipeline integrating latest computer vision techniques.
  - Point cloud reconstruction + instance segmentation (Mask R-CNN & Detectors) + pose estimation (FCGF-based RANSAC & PVN3D) + pose refinement (ICP) + detection filtering (3D NMS)
- Led a team of 20+ engineers in deploying and maintaining 50+ bin-picking systems for continuous operation without failure.

### Research Scientist

ROBOEYE.AI

Toronto, ON, Canada

Mar. 2020 - Jul. 2021

- Implemented an intuitive vision-driven bin-picking solution leveraging a 6D pose estimation pipeline.
  - Fully automated online model training system using PyTorch, NVIDIA Isaac Sim, OpenCV, and AWS
  - C++ application designed for bin-picking tasks using ROS, Qt5, Protobuf, OpenCV, and PCL
  - Online object detection performance tracking system using AWS, Docker, W&B, and Django

### Research Scientist

MOZILLA - EMERGING TECHNOLOGIES TEAM

Remote

Mar. 2020 - Oct. 2020

- Developed a wake-word detection system for Firefox, Howl 🗣️, publishing a first-author paper at an EMNLP 2020 workshop [4].
- Integrated Howl with Firefox Voice 🗣️ to provide a completely hands-free experience to over 8,000 users.

## Research Scientist Intern

SAMSUNG RESEARCH AMERICA - VISUAL DISPLAY INTELLIGENCE LAB

Mountain View, CA, USA

Apr. 2019 - Aug. 2019

- Invented a novel co-clustering algorithm leveraging GANs, resulting in a first-author paper at ICME 2021 [3] and two patents [10, 11].
  - Proposed combining two modality-specific InfoGANs to maximize mutual information between the modalities, identifying unique clusters within dual-modal data that are challenging to detect from a single-modality perspective.
  - Integrated into a recommendation system that captures complex relationships between TV programs and viewers.

## Graduate Student Researcher

UNIVERSITY OF WATERLOO - DATA SYSTEMS GROUP, ADVISOR: PROF. JIMMY LIN

Waterloo, ON, Canada

Sep. 2018 - Dec. 2019

- Honkling: Personalized Keyword Spotting System  - two first-author papers at EMNLP 2019 [7] and IUI 2019 [9]
  - Implemented keyword spotting with convolutional neural networks in pure JavaScript that runs in any standards-compliant browser.
  - Applied on-device fine-tuning with network slimming for accent adaptation and evaluated its efficiency in a browser environment.
- Efficient Domain Adaptation of Language Models
  - Investigated the effects of freezing various BERT layers for multi-task and multilingual tasks - a first-author manuscript [12].
  - Developed memory- and latency-efficient inference techniques for BERT - a paper at ACL 2020 [6].

## Undergraduate Research Assistant

UNIVERSITY OF WATERLOO - ADVISOR: PROF. JIMMY LIN

Waterloo, ON, Canada

May. 2018 - Aug. 2018

- Studied the suitability of JavaScript as an environment for deep learning execution - a first-author manuscript [13].

## Software Engineer Intern

META (FACEBOOK) - DYNAMIC ADS INFRASTRUCTURE

Menlo Park, CA, USA

Jan. 2018 - Apr. 2018

- Applied KNN algorithms on product-level and user-level embeddings to enhance the quality of personalized advertisements.
- Redesigned the advertisements selection pipeline to retrieve user embeddings at an earlier stage, reducing loading time by 7%.

## Undergraduate Research Assistant

UNIVERSITY OF WATERLOO - ADVISOR: PROF. KEN SALEM

Waterloo, ON, Canada

Sep. 2017 - Dec. 2017

- Implemented an RDD usage report generator for Spark  and analyzed the impact of caching replacement policies on performance.

## Undergraduate Research Assistant

UNIVERSITY OF WATERLOO - ADVISOR: PROF. KHUZAIMA DAUDJEE

Waterloo, ON, Canada

Sep. 2017 - Dec. 2017

- Analyzed latency and throughputs of Apache Storm and Spark Streaming; benchmarked against TPCx-IoT specifications.
  - Empirically validated the maximum throughput Apache Storm can handle: 500K tuples per second with a live connection to Redis.

## Software Engineer Intern

UBER - COMPLEX DATA PROCESSING / SPARK TEAM

Palo Alto, CA, USA

May. 2017 - Aug. 2017

- Integrated TensorFlowOnSpark on Uber's infrastructure and evaluated its stability.
- Transformed MLlib pipeline into a Spark job with TensorFlow; reduced training time from 33 to 3 hours.

## Software Engineer Intern

ZYNGA INC - CENTRAL TECHNOLOGY ORGANIZATION

Toronto, ON, Canada

Aug. 2016 - Dec. 2016

- Developed a new architecture for the internal search system.
  - Improved data integrity led to 30% increase in search usage (Amazon Elasticsearch, Amazon Kinesis Streams and Amazon SQS).

## Software Engineer Intern

SAP - EMERGING TECHNOLOGIES TEAM

Waterloo, ON, Canada

Jan. 2016 - Apr. 2016

- Designed and developed a distributed SQLA back-end system with support for the OData protocol.
- Integrated Robot framework, an automated testing tool, to reduce QA cycle from 3 days to 4 hours.

## Software Engineer Intern

MOZZAZ CORPORATION

Waterloo, ON, Canada

May. 2015 - Aug. 2015

- Developed a cross-platform web application using Cordova and Angular.js; performed back-end development with C#.

# Publications and Patents

\* equal contribution

## PUBLICATIONS

- [1] Hwayeon Danielle Shin\*, **Jaejun Lee\***, Federica Guccini. Mining Public Voices: Analyzing Suicide Related Thoughts and Behaviors in YouTube Videos and Comments Using Topic Modeling. *FHLIP*, 2025
- [2] Tomasz Palczewski\*, **Jaejun Lee\***, Lenin Mookiah\*. Production-Ready Applied Deep Learning. *Packt Publishing*, ISBN: 9781803238050, 1803238054, 2022
- [3] **Jaejun Lee**, Hyun Chul Lee, Tomasz Palczewski. CI-GAN: Co-Clustering By Information Maximizing Generative Adversarial Networks. *ICME*, 2021
- [4] Raphael Tang\*, **Jaejun Lee\***, Afsaneh Razi, Julia Cambre, Ian Bicking, Jofish Kaye, Jimmy Lin. Howl: A Deployed, Open-Source Wake Word Detection System. *EMNLP-NLPOSS*, 2020
- [5] Raphael Tang, **Jaejun Lee**, Ji Xin, Xinyu Liu, Yaoliang Yu, Jimmy Lin. Showing Your Work Doesn't Always Work. *ACL*, 2020
- [6] Ji Xin, Raphael Tang, **Jaejun Lee**, Yaoliang Yu, Jimmy Lin. DeeBERT: Dynamic Early Exiting for Accelerating BERT Inference. *ACL*, 2020

- [ 7 ] **Jaejun Lee**, Raphael Tang, Jimmy Lin. Honkling: In-Browser Personalization for Ubiquitous Keyword Spotting. *EMNLP-IJCNLP*, 2019
- [ 8 ] Ryan Clancy, **Jaejun Lee**, Zeynep Akkalyoncu Yilmaz, Jimmy Lin. Information Retrieval Meets Scalable Text Analytics: Solr Integration with Spark. *SIGIR*, 2019
- [ 9 ] **Jaejun Lee**, Raphael Tang, Jimmy Lin. Universal Voice-Enabled User Interfaces using JavaScript. *IUI*, 2019

## PATENTS

- [ 10 ] **Jaejun Lee**, Hyun Chul Lee, Tomasz Palczewski. Co-Informatic Generative Adversarial Networks for Efficient Data Co-Clustering. *International Patent*, Pub. WO/2021/066530, 2021
- [ 11 ] **Jaejun Lee**, Hyun Chul Lee, Tomasz Palczewski. Co-Informatic Generative Adversarial Networks for Efficient Data Co-Clustering. *US Patent*, Pub. 20210097372, 2021

## MANUSCRIPTS

- [ 12 ] **Jaejun Lee**, Raphael Tang, Jimmy Lin. What Would Elsa Do? Freezing Layers During Transformer Fine-Tuning. *arXiv: 1911.03090*, 2019
- [ 13 ] **Jaejun Lee**, Raphael Tang, Jimmy Lin. JavaScript Convolutional Neural Networks for Keyword Spotting in the Browser: An Experimental Analysis. *arXiv: 1810.12859*, 2018

## UNDER REVIEW

- [ 14 ] **Jaejun Lee**, Aristide Tossou, Sandra Wang, Dibyendu Mukherjee. LAM: Few-Shot Random Keypoint Detection in Any Scene through Open-set Object Segmentation. *TAI*

## Presentation

---

### 2025 Future of Health Leadership, Informatics and Policy (FHLIP) Conference

*Toronto, ON, Canada*

ORAL PRESENTATION

*Feb., 2025*

- Mining Public Voices: Analyzing Suicide Related Thoughts and Behaviors in YouTube Videos and Comments Using Topic Modeling.

### Epson's Global Information Sharing Meeting

*Virtual*

ORAL PRESENTATION

*Dec., 2024*

- Leveraging Large Vision-Language Models for Explainable Semantic Concept Learning.

### 2024 American Medical Informatics Association (AMIA) Annual Symposium

*San Francisco, CA, USA*

POSTER PRESENTATION

*Nov., 2024*

- Analyzing YouTube Videos on Suicide-Related Thoughts and Behaviours: A Study Using Topic Modeling and Discourse Analysis.

### 2024 Institute of Health Policy, Management and Evaluation (IHPME) Research & Impact Day

*Toronto, ON, Canada*

ORAL PRESENTATION

*Apr., 2024*

- Analyzing YouTube Videos on Suicide-Related Thoughts and Behaviours: A Study Using Topic Modeling and Discourse Analysis.

### Epson's Global Information Sharing Meeting

*Virtual*

ORAL PRESENTATION

*Apr., 2024*

- Advances in Large Vision and Language Models Driven by Prompt Engineering for Efficient Domain Adaptation.

### 2024 Annual Meeting of the Society for Digital Mental Health

*Virtual*

POSTER PRESENTATION

*Apr., 2024*

- Analyzing YouTube Videos on Suicide-Related Thoughts and Behaviours: A Study Using Topic Modeling and Discourse Analysis.

### Epson's Canadian Information Sharing Meeting

*Markham, ON, Canada*

ORAL PRESENTATION

*Nov., 2023*

- Enhancing 2D Object Detection Efficiency through Prompt Engineering.

### 2nd Workshop for Natural Language Processing Open Source Software (NLP-OSS)

*Virtual*

ORAL PRESENTATION

*Nov., 2020*

- Howl: A Deployed, Open-Source Wake Word Detection System.

### 2019 Conference on Empirical Methods in Natural Language Processing (EMNLP) and 9th International Joint Conference on Natural Language Processing (IJCNLP)

*Hong Kong, China*

POSTER PRESENTATION

*Nov., 2019*

- Honkling: In-Browser Personalization for Ubiquitous Keyword Spotting.

### 24th International Conference on Intelligent User Interfaces (IUI)

*Los Angeles, CA, USA*

POSTER PRESENTATION

*Mar., 2019*

- Universal Voice-Enabled User Interfaces using JavaScript.

# Professional Development

---

## TEACHING ASSISTANTSHIP

### **CS 452/652 – Real-time Programming**

UNIVERSITY OF WATERLOO, INSTRUCTED BY PROF. BILL COWAN

*Waterloo, ON, Canada*

*Fall 2019*

### **CS 480/680 – Introduction to Machine Learning**

UNIVERSITY OF WATERLOO, INSTRUCTED BY PROF. EDITH LAW

*Waterloo, ON, Canada*

*Winter 2019*

### **CS 451/651 – Data Intensive Distributed Computing**

UNIVERSITY OF WATERLOO, INSTRUCTED BY PROF. JIMMY LIN

*Waterloo, ON, Canada*

*Fall 2018*

- Led weekly discussion sessions consisting of 10~20 students and held office hours each week to assist those who needed additional support.
- Graded assignments and exams, and conducted exam preparation sessions.

## ACADEMIC & PROFESSIONAL MENTORSHIP

### **Allen Tao**

RESEARCH INTERN

*Markham, ON, Canada*

*May. 2024 - Present*

- Mentoring a research project on few-shot object and attribute detection.

### **Sandra Wang**

RESEARCH INTERN

*Markham, ON, Canada*

*Mar. 2023 - May. 2024*

- Mentored research projects on 3D scene understanding and few-shot keypoint detection (one paper under review for TAI [14]).

### **Xinyu (Mavis) Liu**

UNDERGRADUATE RESEARCH ASSISTANT

*Waterloo, ON, Canada*

*Jan. 2018 - Aug. 2018*

- Mentored a research project on keyword spotting.