Dhaivat Bhatt

- dhaivat1994@gmail.com - +1(603)-231-4114 - Linkedin
- Google Scholar - Boston, MA - dhaivat1729.github.io

Professional Summary

Experienced AI Researcher with 8 years of expertise in driving AI/ML research in NLP and Multimodal Learning, including 5 years in academia and 3 in industry. Authored over 10 papers for premier conferences like ICML, ICRA, and ECCV, adept at spearheading collaborations that advance practical implementations of cutting-edge technologies.

Skills

- **Programming Languages:** Python, Bash Scripting, Java, C++
- **Python Frameworks and Tools:** PyTorch, transformers, AllenNLP, Seaborn, Matplotlib, Langchain, Tensort-LLM
- **Framework and Tools:** Git, GitHub, AWS, Torchscript, Cursor
- **Relevant Coursework:** Deep learning, Computer vision, Statistical methods in AI, Probabilistic graphical models.

EDUCATION BACKGROUND

Mila - Quebec AI Institute, University of Montreal — **Montreal, Canada**
Research masters - Artificial intelligence — *Sept. 2019 - Aug 2021*

Robotics research center, IIIT Hyderabad — **Hyderabad, India**
MS by research, Computer science and engineering — *Aug. 2016 - Nov. 2018*

BITS Pilani Hyderabad campus — **Hyderabad, India**
BE(Hons.), Electronics and Instrumentation — *Aug. 2012 - Jun. 2016*

PROFESSIONAL EXPERIENCE

Samsung AI Center — **Toronto, Canada**
Deep Learning Research Engineer, Multimodal Learning Group — *September 2021 - September 2024*

- **LLM Agent for phone gallery (Tech used: HF Transformers, Langchain):** Developed an AI agent for phone galleries to assist users in finding and editing media and recording personalized data. Implemented using the ReAct framework on the Llama7B model and integrated 3 different tools for enhanced user experience.
- **Multimodal prompt-tuning of BLIP2 (Tech used: Pytorch, HF Transformers):** *Led* the development and fine-tuning of BLIP2 using prompt-tuning to incorporate non-visual information (names, places, events) into image representations, enhancing image retrieval systems with personalized queries. Resulted in a 20% improvement in retrieval accuracy.
- **Image retrieval evaluation framework (Tech used: Python, Pytorch, Transformers):** *Designed and led* development of a comprehensive evaluation framework for text-based image retrieval systems, adopted across all Samsung AI centers, which standardized benchmarking for over 25 different models.
- **Edge AI Deployment (Tech used: Pytorch, Torchscript, Android Studio):** *Spearheaded* the deployment of CLIP/ALBEF models for edge inference, creating a trace-friendly implementation that enhanced mobile app performance by 25% on Samsung Galaxy S23-Ultra.
- **End-to-end graph parsing architecture (Tech used: Transformers, AllenNLP, Accelerate):** *Designed, ideated and implemented* a differentiable recipe-to-flow graph parsing system, improving F1 scores by 6 points through semi-supervised learning. Highlighted in ECCV 2022 and accepted for full presentation at LREC-COLING 2024. Code available here.
- **Custom LLM Development for QA (Tech used: PyTorch, PeFT, Transformers):** *Led* a project to create a domain-specific LLM for question-answering in the cooking sector, using LoRA with 100 data points to fine-tune Llama-7B on a consumer-grade GPU.
- **QA Dataset Development (Tech used: Prompt Engineering (Llama-13b), Transformers):** *Co-led* a team in constructing an extensive QA dataset with over 12,000 pairs, modeled after SquadV2. The end-to-end QA system established using this dataset became the baseline for future projects.

University of Montreal — **Montreal, Canada**
Visiting researcher, MILA - Quebec AI institute — *November 2018 - July 2019*

- Implemented and tested Deep active localization[link] on a real robot(turtlebot). Successfully ported model trained in simulation to real world setup
- Trained and tested Sparseconvnet models to perform road segmentation in a pointcloud data for Maplite[link]

SELECTED PUBLICATIONS

Organized by theme: *Natural Language Processing (NLP)/Multimodal learning* and *Computer Vision (CV)/Robotics*.

Natural Language Processing (NLP) / Multimodal learning

- *End-to-end Parsing of Procedural Text into Flow Graphs* [LREC-COLING 2024]: **Dhaivat Bhatt***, Ahmad Pourihosseini*, Federico Fancellu, Afsaneh Fazly
- *Flow Graph to Video Grounding for Weakly-Supervised Multi-step Localization* [ECCV 2022 (**oral**)]: Nikita Dvornik, Isma Hajdi, Hai Pham, **Dhaivat Bhatt**, Brais Martinez, Afsaneh Fazly, Allan D. Jepson
- *Visual Semantic Parsing: From Images to Abstract Meaning Representation* [CoNLL 2022]: Mohamed Ashraf Abdelsalam, Zhan Shi, Federico Fancellu, Kalliopi Basioti, **Dhaivat Bhatt**, Vladimir Pavlovic, Afsaneh Fazly

Computer Vision (CV) / Robotics

- *f-Cal: Aleatoric uncertainty quantification for robot perception via calibrated neural regression* [ICRA 2022]: **Dhaivat Bhatt***, Kaustubh Mani*, Dishank Bansal, Krishna Murthy Jatavallabhula, Hanju Lee, Liam Paull
- *MapLite: Autonomous intersection navigation without detailed prior maps* [RAL + ICRA 2020]: Teddy Ort, Krishna Murthy, Rohan Banerjee, Sai Krishna Gottipati, **Dhaivat Bhatt**, Igor Gilitschenski, Liam Paull, Daniela Rus
- *Probabilistic object detection: Strengths, Weaknesses, and Opportunities* [ICML AIAD 2020 Workshop]: **Dhaivat Bhatt***, Dishank Bansal*, Gunshi Gupta*, Hanju Lee, Krishna Murthy Jatavallabhula, Liam Paull
- *Probabilistic obstacle avoidance and object following: An overlap of Gaussians approach* [RO-MAN 2019]: **Dhaivat Bhatt***, Akash Garg*, Bharath Gopalakrishnan, K. Madhava Krishna
- *Have I reached the intersection: A deep learning-based approach for intersection detection from monocular cameras* [IROS 2017]: **Dhaivat Bhatt***, Danish Sodhi*, Arghya Pal, Vineeth Balasubramanian, Madhava Krishna

Reviewer: ICML, CVPR, ICRA, RAL, RO-MAN, LREC-COLING

ACADEMIC PROJECTS

Out of Distribution Detection in Object Detection **Montreal, Quebec**
Prof. Liam Paull, Denso Corporation *Jan 2020 - August 2020*
- Analyzed and critiqued OOD detection methods within object detection frameworks, highlighting significant challenges with the background class problem.

Principled Evaluation of Probabilistic Object Detectors **Montreal, Quebec**
Prof. Liam Paull, MITACS Research Training *July 2020 - November 2020*
- Developed Mahalanobis distance-based criteria integrated into evaluation metrics like mAP, PDQ, and LRP for assessing probabilistic object detectors.

Incremental Learning of Object Detector Through Knowledge Distillation **Montreal, Quebec**
Continual Learning - Course Project *Jan 2020 - April 2020*
- Implemented a pipeline for incremental learning in Detectron2, addressing catastrophic forgetting through knowledge distillation.

Evaluating Robustness of Generative Classifiers Against Adversarial Examples **Montreal, Quebec**
Probabilistic Graphical Models - Course Project *Sept 2019 - Dec 2020*
- Developed a generative classifier resistant to adversarial attacks, validating its robustness under adversarial conditions.

AWARDS AND HONORS

- **May 2020:** Best Paper Award at RAL 2019, ICRA 2020 for the paper "Maplite."
- **Sept 2019:** Awarded a fully funded Research Masters position at MILA - Quebec AI Institute.
- **Aug 2019:** Received a Type C Scholarship, exempting international fees at the University of Montreal ($7,179 CAD per trimester).
- **June 2020:** MITACS Research Training Fellowship, valued at $6,000 CAD.
- **Aug 2017:** Microsoft Research Travel Grant for IROS 2017, totaling 70,000 INR (approximately $1,000 USD in 2017).
- **Jul 2017:** RAS Travel Grant to partially cover expenses for IROS 2017, amounting to $694 USD.
- **Aug 2016:** Research Fellowship at IIIT Hyderabad, covering tuition and living expenses (approximately 350,000 INR or $5,000 USD in 2017).