

# Yash Mukund Kant

University of Toronto

Email : [ysh.kant@gmail.com](mailto:ysh.kant@gmail.com)

Web: [yashkant.github.io](https://yashkant.github.io)

## EDUCATION

---

- **University of Toronto** Ontario, Canada  
*Ph.D. in Computer Science; Advisor: Igor Gilitschenski* *September 2021 – Present*
- **Indian Institute of Technology, Roorkee** Uttarakhand, India  
*B.Tech in Electrical Engineering* *July 2015 – May 2019*

## WORK EXPERIENCE

---

- **Doctoral Researcher** Los Angeles/Toronto  
*Snap Research* *May 2022-Dec 2023*
  - Built and trained 3D-aware Text-to-Multiview Diffusion models; published at CVPR, 2024.
  - Built and trained 3D Diffusion models for Novel View Synthesis; published at SIGGRAPH Asia, 2023.
  - Built a fast and data-driven human animation technique; published at CVPR, 2023.
- **Visiting Research Scholar** Atlanta, GA  
*Georgia Institute of Technology* *September 2019 - August 2021*
  - Built a benchmark to evaluate in LLM-based embodied AI agents; published at ECCV, 2022.
  - Built and trained VQA models that can read and are robust; published at ICCV, 2021 and ECCV, 2020.
  - Runner Up of **TextVQA Challenge** at CVPR, 2020.
  - Organized the **VQA Workshop** and the **TextVQA**, **TextCaps**, and **VQA** challenges at CVPR, 2021.
- **Visiting Research Scholar** Singapore  
*National University of Singapore* *August 2018 - January 2019*
  - **Automated Machine Learning:**
    - \* Designed experiments to search for low memory consuming CNN architectures using Neural Architecture Search.
    - \* Implemented quantization and parameter sharing across the child models to optimize the search procedure.
  - **Resource Efficient Machine Learning:**
    - \* Explored ways to lower the compute of inference and training in ML models with quantization.
- **Software Development Intern** Bengaluru, India  
*Microsoft IDC, Bengaluru* *May 2018 - July 2018*
  - Developed a tool for to automate the creation of environments in Azure Data Lake Store.
  - Built a framework to extract configurations and detect malformed expressions in the Azure codebase.
  - Given a **full-time return offer** for successful completion of the internship project.
- **Winter Research Intern** Singapore  
*National University of Singapore* *November 2017 - January 2018*
  - **Adversarial Machine Learning:**
    - \* Built experiments on adversarial machine learning methods like DeepFool and Fast Gradient Signed Method.
  - **Privacy in Machine Learning:**
    - \* Implemented Membership Inference Attack on ML models under a white-box setting.
- **Software Development Intern** Bengaluru, India  
*Qureto Labs* *May 2017 - July 2017*
  - Developed an android app to handle end-to-end API calls and built UI layouts in XML.
  - Launched the first version a week prior to the proposed deadline, and patched memory issues in the second version.
- **Android Development Intern** Reliance Headquarters, Navi Mumbai  
*Sizzle Labs* *November 2016 - February 2017*
  - Integrated Google Physical Web services in an android application to handle connections with nearby Bluetooth Low Energy beacons.

## TEACHING ASSISTANTSHIPS

---

- **CSC478: Robotics Perception and 3D Vision, Winter 2022**  
*University of Toronto*  
Instructor: **Igor Gilitschenski**  
*Jan 2022 - May 2022*
- **CSC477: Introduction to Mobile Robotics, Fall 2021**  
*University of Toronto*  
Instructor: **Florian Shkurti**  
*August 2021 - December 2021*

## PUBLICATIONS

---

- **SPAD: Spatially Aware Multiview Diffusers**  
**Yash Kant**, Ziyi Wu, Michael Vasilkovsky, Gordon Qian, Jian Ren, Riza Alp Guler, Bernard Ghanem, Sergey Tulyakov\*, Igor Gilitschenski\*, Aliaksandr Siarohin\*  
CVPR, 2024
- **iNVS: Repurposing Diffusion Inpainters for Novel View Synthesis**  
**Yash Kant**, Aliaksandr Siarohin, Michael Vasilkovsky, Riza Alp Guler, Jian Ren, Sergey Tulyakov, Igor Gilitschenski  
Accepted at SIGGRAPH Asia, 2023
- **AvatarOne: Monocular 3D Human Animation**  
Akash Karthikeyan, Robert Ren, **Yash Kant**, Igor Gilitschenski  
Accepted at WACV, 2024
- **Invertible Neural Skinning**  
**Yash Kant**, Aliaksandr Siarohin, Riza Alp Guler, Menglei Chai, Jian Ren, Sergey Tulyakov, Igor Gilitschenski  
Accepted at CVPR, 2023
- **LaTeRF: Label and Text Driven Object Radiance Fields**  
Ashkan Mirzaei, **Yash Kant**, Jonathan Kelly, and Igor Gilitschenski  
Accepted at ECCV, 2022
- **Housekeep: Tidying Virtual Households using Commonsense Reasoning**  
**Yash Kant**, Arun Ramachandran, Sriram Yenamandra, Igor Gilitschenski, Dhruv Batra, Andrew Szot\*, and Harsh Agrawal\*  
Accepted at ECCV, 2022
- **Contrast and Classify: Alternate Training for Robust VQA**  
**Yash Kant**, Abhinav Moudgil, Dhruv Batra, Devi Parikh, Harsh Agrawal.  
Accepted ICCV, 2021
- **Spatially Aware Multimodal Transformers for TextVQA**  
**Yash Kant**, Dhruv Batra, Peter Anderson, Alex Schwing, Devi Parikh, Jiasen Lu, Harsh Agrawal.  
Accepted at ECCV, 2020

## PRE-PRINTS

---

- **Virtual Pets: Animatable Animal Generation in 3D Scenes**  
Yen-Chi Cheng, Chieh Hubert Lin, Chaoyang Wang, **Yash Kant**, Sergey Tulyakov, Alexander Schwing, Liangyan Gui, Hsin-Ying Lee  
In Submission, 2024
- **AToM: Amortized Text-to-Mesh using 2D Diffusion**  
Guocheng Qian, Junli Cao, Aliaksandr Siarohin, **Yash Kant**, Chaoyang Wang, Michael Vasilkovsky, Hsin-Ying Lee, Yuwei Fang, Ivan Skorokhodov, Peiye Zhuang, Igor Gilitschenski, Jian Ren, Bernard Ghanem, Kfir Aberman, Sergey Tulyakov  
In Submission, 2024

## PROJECTS

---

- **Automated YouDescribe** September 2019 - Present  
*Collaborators: Ilmi Yoon and Abhishek Das*  
<https://youdescribe.org/>
  - YouDescribe is a utility tool for visually impaired, which tags and narrates online videos from YouTube.
  - Helping to automate tagging of videos with descriptions using Visual Chatbot and Image Captioning models.
- **Visual Dialog Chatbot** May 2019 - July 2019  
<https://demo.visualdialog.org/>
  - Ported Visual Dialog Chatbot Demo hosted on Cloud-CV from Lua-Torch backend to Pytorch.
  - Trained better Captioning and Visual Dialog models and implemented Beam-Search decoder.
- **ICLR Reproducibility Challenge, 2019** December 2018 - January 2019  
*arXiv: <https://arxiv.org/abs/1901.09517>*  
<https://github.com/yashkant/Padam-Tensorflow>
  - Reproduced the experiments in the ICLR 2019 submission **PADAM** in Tensorflow.
  - Exposed shortcomings of the work and proposed future directions of study.
- **Complement Objective Training for Pythia** February 2019 - April 2019  
*Report: <https://bit.ly/2XnOzZI>*  
<https://github.com/facebookresearch/pythia/pull/32>
  - Added a new training scheme **Complement Objective Training** to **Pythia** (Facebook's open-source framework for vision and language).
  - Analyzed non-convergence issues of Complement Objective Training with Pythia.
- **Quantized Neural Networks** August 2018 – September 2018  
<https://github.com/yashkant/Quantized-Nets>
  - Designed and conducted experiments on Binarized, Ternarized and N-bit Quantized Neural Networks in Tensorflow.
  - Performed a comparative study of quantization schemes and activation functions vs. compute needed for inference.
- **Progressive Neural Architecture Search for Secure Machine Learning** July 2018 – August 2018  
<https://github.com/yashkant/pnas-binary-nets>
  - Designed and added a new penalty function proportional to latency of communication of the MLaaS protocol.
  - Quantized the search space of **Progressive Neural Architecture Search** with binary networks.

## POSITION OF RESPONSIBILITIES

---

- **Joint Secretary** *May 2017 - May 2018*  
*Artificial Intelligence and Electronics Society, IIT Roorkee*
  - Mentored three teams of four members each working on AI and Internet-of-Things related projects.
- **Core Team Member** *January 2017 - May 2017*  
*Institute Technical Council, IIT Roorkee*
  - The group oversees technical reforms in the Institute, proposed circulation of RFID based ID cards.

## REFERENCES

---

**Igor Gilitschenski**  
DCS, University of Toronto  
Research Scientist, TRI  
[igor@gilitschenski.org](mailto:igor@gilitschenski.org)

**Devi Parikh**  
CoC, Georgia Tech  
Research Director, FAIR  
[parikh@gatech.edu](mailto:parikh@gatech.edu)

**Dhruv Batra**  
CoC, Georgia Tech  
Research Director, FAIR  
[dbatra@gatech.edu](mailto:dbatra@gatech.edu)