Vikram Voleti

PhD candidate at Mila; former Research Intern at Google, Unity, Meta ; 4+ years of work experience

voletiv.github.io | vikram.voleti@gmail.com | Google Scholar | LinkedIn

Research

Deep learning for image, video, 3D: expert at machine learning research and development; experienced in leading multiple projects collaborating with international partners in industry and academia.

Projects include:
- Score-based denoising diffusion models for video [1], deriving non-isotropic covariance [2]
- Image generation using normalizing flows [4][9]; video generation using Neural ODEs [14], GANs [16][17]
- 3D human pose estimation and inverse kinematics [3], 3D object generation using NeRFs, diffusion
- Contributed to projects on 4D generation, simulation [10], fairness/uncertainty [5][12], federated learning [6]

Education

Mila, University of Montreal, Canada — ***Sep 2018 - present (Aug 2023)***

Ph.D. in Computer Science — *Supervisor*: Prof. Christopher Pal

Indian Institute of Technology (IIT), Kharagpur, India — ***2009 - 2014***

Dual Degree (**B.Tech. (Honours) + M.Tech.**) in Electrical Engineering with Master's specialization in Instrumentation and Signal Processing

Research Internships during PhD

Meta (formerly **Facebook**), Menlo Park, USA — ***Aug-Dec 2022***

Team: AI for Metaverse (AI4RL); *Supervisors*: Dr. Yashar Mehdad, Dr. Barlas Oguz
- Research and development of solutions for text to 3D object generation using diffusion models, NeRF
- Led project in collaboration with international teams, applied research to virtual reality product

Unity Technologies, Montreal, Canada (MITACS Research Intern) — ***Oct 2021 - Aug 2022***

Team: Deep Pose, Unity Labs; *Supervisor*: Dr. Boris Oreshkin
- 3D human pose estimation and inverse kinematics from videos, published at SIGGRAPH Asia [3]
- Led project on AI-assisted animation workflows, contributed to product pipeline with code, demos

Google, Mountain View, USA — ***Sep-Dec 2019***

Team: Google AI Perception; *Supervisors*: Dr. Bryan Seybold, Dr. Sourish Chaudhuri
- Research on multimodal semi-supervised Active Speaker Detection in videos

Work Experience

IIIT Hyderabad, India — Research Fellow; *Supervisor*: Prof. C. V. Jawahar — ***May 2017 - Aug 2018***
- Synthesized educational videos in regional Indian languages by generating lips from audio
- Developed automated pipeline to create large-scale audio-video dataset
- Full paper published at ICASSP 2019 [16], short paper published at CVPR 2018 Workshop [17]

GreyOrange Robotics, Gurgaon, India — Image Processing Engineer — ***Feb 2016 - May 2017***
- Developed computer vision solutions for robotic applications in real time for warehouse automation
- Solely responsible for development and testing of code, video processing module, camera drivers, server

Airbus, Bengaluru, India — Associate Engineer — ***Jul 2014 - Feb 2016***
- Avionics software development and integration following standard avionics coding guidelines (DO-178B)

Other Professional Experience

Blue Lion Labs, Canada — AI Advisor — ***Oct 2020 - present***
- Provide technical guidance and mentorship on the design and development of AI/ML systems
- Mentor co-op students and interns, published research papers from work led by them [6][8]

NextAI, Canada — AI Scientist-in-Residence — ***Apr-Sep 2019, Mar-Sep 2020***
- Provided scientific and technical support to start-ups selected in yearly co-hort of NextAI accelerator

Playment, Bengaluru, India — Computer Vision Consultant — ***Jan-Jun 2018***
- Provided technical guidance on semantic segmentation models for autonomous driving

TalentSprint, Hyderabad, India — Mentor, Foundations of AI & ML (inaugural program) — ***Jan-May 2018***
- Designed and delivered tutorials on machine learning, and provided mentorship to industry professionals

Awards

Outstanding Reviewer at CVPR 2021	*May 2021*
Microsoft Diversity Award for Doctoral Research, $6,000	*Dec 2020*
MITACS Accelerate Research Internship, $30,000	*Oct 2020*
University of Montreal entrance scholarship, $37,000	*Sep 2018*
IIIT Hyderabad merit scholarship for summer school, $1,000	*Jul 2017*

Service

Organizer — **ICCV 2021** - Differentiable 3D Vision and Graphics workshop *Feb-Oct 2021*
OWCV 2021 (Canadian Computer Vision workshop), Canada *Feb-Apr 2021*
GRAPHQUON 2020 (Canadian Computer Graphics workshop), Canada *Oct-Dec 2020*

Reviewer — Journal on Computer Vision and Image Understanding, CVPR 2022, ACML 2021, NeurIPS 2021, ICCV 2021, CVPR 2021 (***Outstanding Reviewer***), ICLR 2020, NeurIPS 2020, ICML 2020, NeurIPS 2019, CCAI @ ICLR 2020, CCAI @ NeurIPS 2019, LLD @ ICLR 2019

Skills

C/C++, CUDA, HTML/CSS, Javascript, Jax, Keras, LaTeX, MATLAB, OpenCV, OS X, Python, PyTorch, R, Shell, SLURM, Tensorflow, Ubuntu, Verilog, Windows

Deep learning, computer vision, machine learning, research and development, generative modeling, NeRF, score-based diffusion models, normalizing flows, Neural ODEs, GANs, Transformers, large-scale training, image generation, video prediction, 3D pose estimation, 3D rendering, text-to-image, text-to-3D, text-to-4D

Talks

- "Diffusion models for solving video tasks" — INRIA, France [slides] *Feb 2023*
- "MCVD: Masked Conditional Video Diffusion" — NeurIPS 2022, New Orleans, USA [slides] *Dec 2022*
- "SMPL-IK: Learned Morphology-Aware Inverse Kinematics for AI Driven Artistic Workflows" — SIGGRAPH Asia, Diagu, South Korea [slides, video] *Dec 2022*
- "Normalizing flows" — Learning Representations (course), University of Montreal, Canada *Nov 2022*
- "Score-based Denoising Diffusion Models - a tutorial" — Mila, Canada [slides, video] *Sep 2022*
- "Solving Video Tasks using Denoising Diffusion Models" — Samsung Toronto, Canada [slides] *Aug 2022*
- "MCVD: Masked Conditional Video Diffusion" — Mila, Canada *May 2022*
- "Denoising Diffusion GANs" — Mila, Canada [slides] *Feb 2022*
- "Training GANs by Solving ODEs" — Mila, Canada [slides] *Apr 2021*
- "Score-based Generative Models with SDEs" — Mila, Canada [slides] *Feb 2021*
- "Continuous Normalizing Flows" — Mila, Canada [slides] *Sep 2020*
- "GANs: the story so far" — Summer Symposium on AI Research, India [slides, video] *Jul 2020*
- "A brief tutorial on Neural ODEs" — Mila, Canada [slides, video] *Jul 2020*
- "Mathematics of Neural ODEs" — University of Guelph, Canada [slides] *Apr 2020*
- "Simple Video Generation using Neural ODEs" — IIIT Hyderabad, India [slides] *Jan 2020*
- Tutorial on "GANs" — AI for Social Good Summer Lab, Montreal *May 2019*
- "BigGAN" — Mila, University of Montreal, Canada [slides] *Oct 2018*
- "Image de-fencing using RGB-D data" — MPI Informatics, Saarbrücken, Germany [slides] *Feb 2018*
- "Intuition behind LSTMs" — IIIT Hyderabad, India [slides] *Feb 2018*
- Tutorial on "Back-propagation" — IIIT-Hyderabad, India [slides] *Aug 2017*
- "Mathematics of back-propagation" — GreyOrange Robotics, India [slides] *Feb 2017*

Past Internships

KU Leuven, Belgium — *Supervisor*: Prof. Ingrid Verbauwhede, ESAT ***Summer 2013***
- Designed arithmetic operations using Carry-Free Logic, simulated circuits in Xilinx

IIT Kharagpur, India — *Supervisor*: Prof. Aurobinda Routray, Electrical Engineering ***Summer 2012***
- Made a gesture recognition program in MATLAB using Hidden Markov Models

Imperial College, UK — *Supervisor*: Prof. Peter Cheung, Electrical & Electronics ***Summer 2011***
- Measured the relative power consumption among the LookUp Tables (LUTs) of an FPGA

Thesis Projects

Supervisor: Prof. Rajiv Sahay, Electrical Engineering, IIT Kharagpur, India

Master's thesis — "De-fencing of Images using RGB-D Data" ***2013 - 2014***
- Elimination of fence-like occlusions, and inpainting of images using RGB-D data
- Nominated for Best Project Award among three departments, research work published at ICAPR 2015

Bachelor's thesis — "Identification of Bilabial Lip Closures in Audio and Video" ***2012 - 2013***
- Measurement of synchronization between audio and video using bilabial cues in both modes

Research Papers (Select)

[1] "MCVD: Masked Conditional Video Diffusion for Prediction, Generation, and Interpolation", **V. Voleti**, A. Jolicoeur-Martineau, C. Pal - *NeurIPS 2022* [arXiv]

[2] "Score-based Denoising Diffusion with Non-Isotropic Gaussian Noise Models", **V. Voleti**, C. Pal, A. Oberman - *NeurIPS 2022 Workshop* [arXiv]

[3] "SMPL-IK: Learned Morphology-Aware Inverse Kinematics for AI Driven Artistic Workflows", **V. Voleti**, B. N. Oreshkin, F. Bocquelet, F. G. Harvey, L. Ménard, C. Pal - *SIGGRAPH Asia 2022* [arXiv]

[4] "Multi-Resolution Continuous Normalizing Flows", **V. Voleti**, C. Finlay, A. Oberman, C. Pal - *Submitted to a journal* [arXiv]

[5] "FairCal : Fairness Calibration for Face Verification", T. Salvador, S. Cairns, **V. Voleti**, N. Marshall, A. Oberman - *ICLR 2022* [arXiv]

[6] "Plankton-FL: Exploration of Federated Learning for Privacy-Preserving Training of Deep Neural Networks for Phytoplankton Classification", D. Zhang, **V. Voleti**, A. Wong, J. Deglint - *CVIS 2022 (Oral)*

[7] "Generative Models of Brain Dynamics", M. Ramezanian-Panahi, G. Abrevaya, JC. Gagnon-Audet, **V. Voleti**, I. Rish, G. Dumas - *Frontiers in Artificial Intelligence (journal)* [arXiv]

[8] "Towards Generating Large Synthetic Phytoplankton Datasets for Efficient Monitoring of Harmful Algal Blooms", N. Bamra, **V. Voleti**, A. Wong, J. Deglint - *FSS at AAAI 2022* [arXiv]

[9] "Improving Continuous Normalizing Flows using a Multi-Resolution Framework", **V. Voleti**, C. Finlay, A. Oberman, C. Pal - *ICML 2021 Workshop*

[10] "gradSim: Differentiable simulation for system identification and visuomotor control" , K. M. Jatavallabhula, M. Macklin, F. Golemo, **V. Voleti**, L. Petrini, M. Weiss, B. Considine, J. Parent-Lévesque, K. Xie, K. Erleben, L. Paull, F. Shkurti, D. Nowrouzezahrai, S. Fidler - *ICLR 2021* [arXiv]

[11] "Accounting for Variance in Machine Learning Benchmarks", X. Bouthillier, P. Delaunay, M. Bronzi, A. Trofimov, B. Nichyporuk, J. Szeto, N. Sepah, E. Raff, K. Madan, **V. Voleti**, S. E. Kahou, V. Michalski, D. Serdyuk, T. Arbel, C. Pal, G. Varoquaux, P. Vincent - *MLSys 2021* [arXiv]

[12] "Frustratingly Easy Uncertainty Estimation for Distribution Shift", T. Salvador, **V. Voleti**, A. Iannantuono, A. Oberman - *Preprint* [arXiv]

[13] "Learning to Combine Top-Down and Bottom-Up Signals in RNNs with Attention over Modules", S. Mittal, A. Lamb, A. Goyal, **V. Voleti**, M. Shanahan, G. Lajoie, M. Mozer, Y. Bengio - *ICML 2020* [arXiv]

[14] "Simple Video Generation using Neural ODEs", **V. Voleti**, D. Kanaa, S. E. Kahou, C. Pal - *NeurIPS 2019 Workshop* [arXiv]

[15] "Comparing Normalization in Conditional Computation Tasks", V. Michalski, **V. Voleti**, S. E. Kahou, A. Oritz, P. Vincent, C. Pal, D. Precup - *ICML 2019 Workshop* [arXiv]

[16] "Cross-Language Speech Dependent Lip-Synchronization", **V. Voleti**, A. Jha, V. P. Namboodiri, C. V. Jawahar - *ICASSP 2019* [pdf]

[17] "Lip-Synchronization for Dubbed Instructional Videos", **V. Voleti**, A. Jha, V. P. Namboodiri, C. V. Jawahar - *CVPR 2018 Workshop* (FIVER) [pdf]

[18] "A Multimodal Approach for Image De-fencing and Depth Inpainting", S. Jonna, **V. Voleti**, R. R. Sahay, and M. S. Kankanhalli - *ICAPR* 2015 [pdf, IEEE]

Teaching Experience

University of Montreal, Montreal, Canada — Guest Lecturer ***Nov 2020***
- Representation Learning (IFT 6135) by Prof. Aishwarya Agrawal

University of Montreal, Montreal, Canada — Teaching Assistant ***Sep-Dec 2020***
- Fundamentals of Machine Learning (IFT 6390) by Prof. Ioannis Mitliagkas

Summer Symposium on AI Research, India — Guest Speaker ***Jul 2020***

University of Montreal, Montreal, Canada — Teaching Assistant ***Sep 2019***
- Fundamentals of Machine Learning (IFT 6390) by Prof. Ioannis Mitliagkas

IVADO/Mila Deep Learning School, Montreal, Canada — Teaching Assistant ***Sep 2019***

AI for Social Good Summer Lab, Montreal, Canada — Lecturer ***May 2019***

TalentSprint, Hyderabad, India — Mentor, Foundations of AI & ML (inaugural program) ***Jan-May 2018***
- Designed and presented tutorials on machine learning, and mentored industry professionals