



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](#)

EXPERTISE	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 [13], GANs [15][16] • 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], federated learning [6]		
EDUCATION	 Mila, University of Montreal , Canada <i>Sep 2018 - present (Aug 2023)</i> Ph.D. in Computer Science — <i>Supervisor:</i> Prof. Christopher Pal		
	 Indian Institute of Technology (IIT), Kharagpur , India <i>2009 - 2014</i> Dual Degree (B.Tech. (Honours) + M.Tech.) in Electrical Engineering with Master's specialization in Instrumentation and Signal Processing		
RESEARCH	 Meta (formerly Facebook), Menlo Park, USA <i>Aug-Dec 2022</i>		
INTERNSHIPS	<i>Team:</i> AI for Metaverse (AI4RL); <i>Supervisors:</i> Dr. Yashar Mehdad, Dr. Barlas Oguz		
DURING PHD	<ul style="list-style-type: none">Led the technology development for generating 3D objects, videos from text; diffusion models, NeRFApplied expertise at neural graphics for 3D rendering; implemented hands-on in PyTorchInternational AI team; technology transitioned into a Meta end product, and adopted by other Meta teams		
	 Unity Technologies , Montreal, Canada (MITACS Research Intern) <i>Oct 2021 - Aug 2022</i>		
	<i>Team:</i> Deep Pose, Unity Labs; <i>Supervisor:</i> Dr. Boris Oreshkin		
	<ul style="list-style-type: none">Built AI-assisted animation workflow for user-editable 3D characters; trained novel 3D human pose priorPublished at SIGGRAPH Asia [3], incorporated technology into a Unity product		
	 Google , Mountain View, USA <i>Sep-Dec 2019</i>		
	<i>Team:</i> Google AI Perception; <i>Supervisors:</i> Dr. Bryan Seybold, Dr. Sourish Chaudhuri		
	<ul style="list-style-type: none">Investigated the scope of deep semi-supervised learning for active speaker detection in videoHands-on implementation in TensorFlow; collaborated with TPU team for coding Neural ODE in Jax/Flax		
WORK	IIT Hyderabad , India — Research Fellow; <i>Supervisor:</i> Prof. C. V. Jawahar <i>May 2017 - Aug 2018</i>		
EXPERIENCE	<ul style="list-style-type: none">Synthesized educational videos in regional Indian languages by generating lips from audioDeveloped automated pipeline to create large-scale audio-video datasetFull paper published at ICASSP 2019 [15], short paper published at CVPR 2018 Workshop [16]		
	GreyOrange Robotics , Gurgaon, India — Image Processing Engineer <i>Feb 2016 - May 2017</i>		
	<ul style="list-style-type: none">Developed computer vision solutions for embedded robotics in real time for warehouse automationSolely responsible for code development and testing of video processing module, camera drivers, server		
	Airbus , Bengaluru, India — Associate Engineer <i>Jul 2014 - Feb 2016</i>		
	<ul style="list-style-type: none">Avionics software development following standard avionics coding guidelines (DO-178B)		
	Blue Lion Labs , Canada — AI Advisor <i>Oct 2020 - present</i>		
	<ul style="list-style-type: none">Provide technical guidance and mentorship on the design and development of AI/ML systems		
PROFESSIONAL	<ul style="list-style-type: none">Mentor co-op students and interns, published research papers from work led by them [6][8]		
EXPERIENCE	NextAI , Canada — AI Scientist-in-Residence <i>Apr-Sep 2019, Mar-Sep 2020</i>		
	<ul style="list-style-type: none">Provided scientific and technical support to start-ups selected in yearly co-hort of NextAI accelerator		
	Playment , Bengaluru, India — Computer Vision Consultant <i>Jan-Jun 2018</i>		
	<ul style="list-style-type: none">Provided technical guidance on semantic segmentation models for autonomous driving		
	TalentSprint , Hyderabad, India — Mentor, Foundations of AI & ML (inaugural program) <i>Jan-May 2018</i>		
	<ul style="list-style-type: none">Designed and delivered tutorials on machine learning, and provided mentorship to industry professionals		
AWARDS	Outstanding Reviewer at CVPR 2021	<i>May 2021</i>	
	Microsoft Diversity Award for Doctoral Research, \$6,000	<i>Dec 2020</i>	
	MITACS Accelerate Research Internship, \$30,000	<i>Oct 2020</i>	
	University of Montreal entrance scholarship, \$37,000	<i>Sep 2018</i>	
	IIT Hyderabad merit scholarship for summer school, \$1,000	<i>Jul 2017</i>	

SERVICE	Organizer — ICCV 2021 - Differentiable 3D Vision and Graphics workshop OWCV 2021 (Canadian Computer Vision workshop), Canada GRAPHQUON 2020 (Canadian Computer Graphics workshop), Canada Reviewer — Journal on Computer Vision and Image Understanding, CVPR 2022, ACML 2021, NeurIPS 2021, ICCV 2021, CVPR 2021 (<i>Outstanding Reviewer</i>), ICLR 2020, NeurIPS 2020, ICML 2020, NeurIPS 2019, CCAI @ ICLR 2020, CCAI @ NeurIPS 2019, LLD @ ICLR 2019	Feb-Oct 2021 Feb-Apr 2021 Oct-Dec 2020
TEACHING EXPERIENCE	University of Montreal , Montreal, Canada — Guest Lecturer • Representation Learning (IFT 6135) by Prof. Aishwarya Agrawal University of Montreal , Montreal, Canada — Teaching Assistant • Fundamentals of Machine Learning (IFT 6390) by Prof. Ioannis Mitliagkas Summer Symposium on AI Research , India — Guest Speaker University of Montreal , Montreal, Canada — Teaching Assistant • Fundamentals of Machine Learning (IFT 6390) by Prof. Ioannis Mitliagkas IVADO/Mila Deep Learning School , Montreal, Canada — Teaching Assistant AI for Social Good Summer Lab , Montreal, Canada — Lecturer TalentSprint , Hyderabad, India — Mentor, Foundations of AI & ML (inaugural program) • Designed and presented tutorials on machine learning, and mentored industry professionals	Nov 2020 Sep-Dec 2020 Jul 2020 Sep 2019 Sep 2019 May 2019 Jan-May 2018
PAST INTERNSHIPS	KU Leuven , Belgium — <i>Supervisor</i> : Prof. Ingrid Verbauwhede, ESAT • Designed and implemented carry-free arithmetic operations in Verilog; simulated circuits in Xilinx IIT Kharagpur , India — <i>Supervisor</i> : Prof. Aurobinda Routray, Electrical Engineering • Made a gesture recognition program in MATLAB using Hidden Markov Models Imperial College , UK — <i>Supervisor</i> : Prof. Peter Cheung, Electrical & Electronics • Circuits and Systems Research Group; measured intra-die power variation in sub-nm FPGAs	Summer 2013 Summer 2012 Summer 2011
SKILLS	C/C++, CUDA, HTML/CSS, Javascript, Jax, Keras, L ^A T _E X, 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] • “MCVD: Masked Conditional Video Diffusion” — NeurIPS 2022, New Orleans, USA [slides] • “SMPL-IK: Learned Morphology-Aware Inverse Kinematics for AI Driven Artistic Workflows” — SIGGRAPH Asia, Diagu, South Korea [slides, video] • “Normalizing flows” — Learning Representations (course), University of Montreal, Canada • “Score-based Denoising Diffusion Models - a tutorial” — Mila, Canada [slides, video] • “Solving Video Tasks using Denoising Diffusion Models” — Samsung Toronto, Canada [slides] • “MCVD: Masked Conditional Video Diffusion” — Mila, Canada • “Denoising Diffusion GANs” — Mila, Canada [slides] • “Training GANs by Solving ODEs” — Mila, Canada [slides] • “Score-based Generative Models with SDEs” — Mila, Canada [slides] • “Continuous Normalizing Flows” — Mila, Canada [slides] • “GANs: the story so far” — Summer Symposium on AI Research, India [slides, video] • “A brief tutorial on Neural ODEs” — Mila, Canada [slides, video] • “Mathematics of Neural ODEs” — University of Guelph, Canada [slides] • “Simple Video Generation using Neural ODEs” — IIIT Hyderabad, India [slides] • Tutorial on “GANs” — AI for Social Good Summer Lab, Montreal • “BigGAN” — Mila, University of Montreal, Canada [slides] • “Image de-fencing using RGB-D data” — MPI Informatics, Saarbrücken, Germany [slides] • “Intuition behind LSTMs” — IIIT Hyderabad, India [slides] • Tutorial on “Back-propagation” — IIIT-Hyderabad, India [slides] • “Mathematics of back-propagation” — GreyOrange Robotics, India [slides]	Feb 2023 Dec 2022 Dec 2022 Nov 2022 Sep 2022 Aug 2022 May 2022 Feb 2022 Apr 2021 Feb 2021 Sep 2020 Jul 2020 Jul 2020 Apr 2020 Jan 2020 May 2019 Oct 2018 Feb 2018 Feb 2018 Aug 2017 Feb 2017

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] *NeurIPS 2022* - “MCVD: Masked Conditional Video Diffusion for Prediction, Generation, and Interpolation”, **V. Voleti**, A. Jolicoeur-Martineau, C. Pal [[arXiv](#)]
- [2] *NeurIPS 2022 Workshop* - “Score-based Denoising Diffusion with Non-Isotropic Gaussian Noise Models”, **V. Voleti**, C. Pal, A. Oberman [[arXiv](#)]
- [3] *SIGGRAPH Asia 2022* - “SMPL-IK: Learned Morphology-Aware Inverse Kinematics for AI-Driven Artistic Workflows”, **V. Voleti**, B. N. Oreshkin, F. Bocquet, F. G. Harvey, L. Ménard, C. Pal [[arXiv](#)]
- [4] *Submitted to a journal* - “Multi-Resolution Continuous Normalizing Flows”, **V. Voleti**, C. Finlay, A. Oberman, C. Pal [[arXiv](#)]
- [5] *ICLR 2022* - “FairCal : Fairness Calibration for Face Verification”, T. Salvador, S. Cairns, **V. Voleti**, N. Marshall, A. Oberman [[arXiv](#)]
- [6] *CVIS 2022 (Oral)* - “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
- [7] *Frontiers in Artificial Intelligence (journal)* - “Generative Models of Brain Dynamics”, M. Ramezani-Panahi, G. Abrevaya, J.C. Gagnon-Audet, **V. Voleti**, I. Rish, G. Dumas [[arXiv](#)]
- [8] *FSS at AAAI 2022* - “Towards Generating Large Synthetic Phytoplankton Datasets for Efficient Monitoring of Harmful Algal Blooms”, N. Bamra, **V. Voleti**, A. Wong, J. Deglint [[arXiv](#)]
- [9] *ICML 2021 Workshop* - “Improving Continuous Normalizing Flows using a Multi-Resolution Framework”, **V. Voleti**, C. Finlay, A. Oberman, C. Pal
- [10] *ICLR 2021* - “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 [[arXiv](#)]
- [11] *MLSys 2021* - “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 [[arXiv](#)]
- [12] *ICML 2020* - “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 [[arXiv](#)]
- [13] *NeurIPS 2019 Workshop* - “Simple Video Generation using Neural ODEs”, **V. Voleti**, D. Kanaa, S. E. Kahou, C. Pal [[arXiv](#)]
- [14] *ICML 2019 Workshop* - “Comparing Normalization in Conditional Computation Tasks”, V. Michalski, **V. Voleti**, S. E. Kahou, A. Oritz, P. Vincent, C. Pal, D. Precup [[arXiv](#)]
- [15] *ICASSP 2019* - “Cross-Language Speech Dependent Lip-Synchronization”, **V. Voleti**, A. Jha, V. P. Namboodiri, C. V. Jawahar [[pdf](#)]
- [16] *CVPR 2018 Workshop* - “Lip-Synchronization for Dubbed Instructional Videos”, **V. Voleti**, A. Jha, V. P. Namboodiri, C. V. Jawahar (FIVER) [[pdf](#)]
- [17] *ICAPR 2015* - “A Multimodal Approach for Image De-fencing and Depth Inpainting”, S. Jonna, **V. Voleti**, R. R. Sahay, and M. S. Kankanhalli [[pdf](#), [IEEE](#)]