Vikram Voleti

vikram.voleti@gmail.com

Online

Website: voletiv.github.io | *Google Scholar* | *LinkedIn* | *GitHub*

Education

PhD student, with Prof. Christopher Pal — *Fall 2018 - present*
Mila, University of Montreal, Canada

Dual Degree (B.Tech. (H) + M.Tech.) in Electrical Engineering — *2009 - 2014*
with Master's specialization in Instrumentation and Signal Processing
Indian Institute of Technology (IIT), Kharagpur, India — CGPA: 8.44 / 10

Research Papers

[1] Vikram Voleti, David Kanaa, Samira E. Kahou, Chris Pal, "Simple Video Generation using Neural ODEs" - *NeurIPS 2019 Workshop* (LIRE) [pdf]

[2] Vincent Michalski, Vikram Voleti, Samira E. Kahou, Anthony Oritz, Pascal Vincent, Chris Pal, Doina Precup, "Comparing Normalization in Conditional Computation Tasks" - *ICML 2019 Workshop* [pdf]

[3] Abhishek Jha*, Vikram Voleti*, Vinay P. Namboodiri, C. V. Jawahar, "Cross-Language Speech Dependent Lip-Synchronization" - *ICASSP 2019* [pdf]

[4] Abhishek Jha*, Vikram Voleti*, Vinay P. Namboodiri, C. V. Jawahar, "Lip-Synchronization for Dubbed Instructional Videos" - *CVPR 2018 Workshop* (FIVER) [pdf, url]

[5] V. Voleti, P. Mohan, S. Gupta, J. Iqbal, "Simple Real-Time Pattern Recognition for Industrial Automation" - *Proc. International Conference on Industrial Design Engineering*, 2017 [pdf]

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

Research Experience

Visiting Researcher — *Prof. Graham Taylor*, University of Guelph, Canada — *Dec 2019 - present*

Research Intern — Google, Mountain View, USA — *Sep 2019 - Dec 2019*
Google AI Perception team — Bryan Seybold, Sourish Chaudhuri

- Research on Semi-supervised Active Speaker Detection in videos
- Research on using Switching Non-Linear Dynamical Systems to model speaker activity

Research Fellow — *Prof. C. V. Jawahar*, IIIT Hyderabad, India — *May 2017 - Aug 2018*

- Built a visual speech recognizer (lipreader) to classify spoken words without audio
- Synthesized video in regional Indian languages by generating lips from audio
- Full paper published at ICASSP 2019 [3], short paper published at CVPR Workshop 2018 [4]

Other Experience

Reviewer — ICML 2020, ICLR 2020, CCAI @ ICLR 2020, CCAI @ NeurIPS 2019, NeurIPS 2019

Teaching Assistant — IVADO/Mila Deep Learning School, Montreal, Canada — *Sep 9-13, 2019*

Teaching Assistant — University of Montreal, Montreal, Canada

- Fundamentals of Machine Learning (IFT 6390) — Ioannis Mitliagkas — *Sep 2019*

Scientist in Residence — NextAI (startup accelerator), Montreal, Canada — *Apr 2019 - Aug 2019*

- Consultant for multiple startups on computer vision, deep learning and AI

Consultant, Computer Vision — Playment, Bengaluru, India — *Jan 2018 - Jun 2018*

- Worked on semantic segmentation models for autonomous driving

Mentor, Foundations of AI & ML — TalentSprint, Hyderabad, India — *Jan 2018 - May 2018*
Six months certificate program in collaboration with IIIT Hyderabad, *India*

- Designed and presented tutorials on machine learning, and mentored industry professionals

Work Experience

Image Processing Engineer — GreyOrange Robotics, Gurgaon, India *Feb 2016 - May 2017*

- Developed computer vision module for video processing in real time for warehouse automation
- Research paper [5] based on work is published by ACM at ICIDE 2017

Associate Engineer — Airbus Group India, Bengaluru, India *Jul 2014 - Feb 2016*

- Involved in development and integration of avionics systems for the long-range aircrafts family
- Simulated signal-level modifications to the Flight Warning Computer, adopting standard avionics coding guidelines (DO-178B)

Talks & Other Efforts

- *Jan 2020* - Talk: Simple Video Generation using Neural ODEs at IIIT Hyderabad, India [presentation]
- *May 2019* - Talk: Tutorial on GANs at the AI for Social Good Summer Lab, Montreal
- *Jan 2019* - Code: Released code for Self-Attention GAN in PyTorch, converting from TensorFlow code released by Google Brain [GitHub]
- *Oct 2018* - Talk: "BigGAN - Large Scale GAN Training for High Fidelity Natural Image Synthesis" at Mila, University of Montreal, Canada [presentation]
- *Feb 2018* - Talk: "Image de-fencing using RGB-D data" at Max Planck Insitute for Informatics, Saarbrücken, Germany [presentation]
- *Feb 2018* - Talk: "Intuition behind LSTMs" at IIIT Hyderabad, India [presentation]
- *Aug 2017* - Talk: "Mathematics of back-propagation in multi-layer perceptrons" at GreyOrange Robotics, India, and at IIIT-Hyderabad, India [tutorial]
- Attended summer schools on Computer Vision and Machine Learning at IIIT-Hyderabad in 2017
  - Stood 3rd in Computer Vision Summer School out of 120+ participants, rewarded full fee waiver
  - Stood 4th in Machine Learning Summer School out of 120+ participants, rewarded full fee waiver
- Won the SMS Classification challenge, participated in the Video Action Recognition challenge in the 2017 Hack2Innovate hackathon in Bangalore, India
- Qualified JEE 2009 by IIT at 99.7 percentile, with All India Rank of 1330 (out of 384,977)

Thesis Projects

Masters thesis — "De-fencing of Images using RGB-D Data" *2013 - 2014*
IIT Kharagpur, India — *Prof. Rajiv Sahay, Department of Electrical Engineering*

- Elimination of fence-like occlusions, and inpainting of images using RGB-D data
- Nominated for Best M.Tech. Project Award among three departments (Electrical, Electronics, CS)
- Research paper [6] based on work is published in the proceedings of ICAPR 2015
- Links — GitHub repository containing thesis, presentation, code files, and results

Bachelors thesis — "Identification of Bilabial Lip Closures in Audio and Video" *2012 - 2013*
IIT Kharagpur, India — *Prof. Rajiv Sahay, Department of Electrical Engineering*

- Measurement of synchronization between audio and video using bilabial cues in both modes
- Links — GitHub repository containing thesis, presentation, code files, and results

---

Past Research Internships

KU Leuven, Belgium — *Prof. Ingrid Verbauwhede, ESAT* *Summer 2013*

- Designed arithmetic operations using Carry-Free Logic, simulated circuits in Xilinx

IIT Kharagpur, India — *Prof. Aurobinda Routray, Electrical Engineering* *Summer 2012*

- Made a gesture recognition program in MATLAB using Hidden Markov Models

Imperial College, London, UK — *Prof. Peter Cheung, Electrical & Electronics* *Summer 2011*

- Measured the relative power consumption among the LookUp Tables (LUTs) of an FPGA

Skills

C/C++, CUDA, HTML/CSS, Javascript, Keras, MATLAB, OpenCV, Python, PyTorch, Shell, Tensorflow