Vikram Voleti

Online — *Website*: voletiv.github.io *GitHub:* github.com/voletiv *LinkedIn*: Vikram Voleti

Education

PhD student at Mila, with Prof. Christopher Pal — *Fall 2018 - present*
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] Vincent Michalski, Vikram Voleti, Samira E. Kahou, Anthony Oritz, Pascal Vincent, Chris Pal, Doina Precup, "Comparing Normalization in Conditional Computation Tasks" in *ICML 2019 Workshop* [pdf]

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

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

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

[5] S. Jonna, V. S. Voleti, R. R. Sahay, and M. S. Kankanhalli, "A Multimodal Approach for Image De-fencing and Depth Inpainting," in *ICAPR*, 2015, pp. 1—6 [pdf, IEEE]

Current Projects

- Studying the effect of normalization in conditional image generation [1]
- Visual reasoning via language grounding: integrating NLP into GANs for Visual QA
- Other projects: deep generative models for 3D, video prediction

Research Experience

Research Fellow — *Applied Research Lab* — *May 2017 - August 2018*
International Institute of Information Technology (IIIT) - Hyderabad, India
Prof. C. V. Jawahar, Centre for Visual Information Technology, IIIT-Hyderabad

PROJECT: **Video Translation**
- Synthesized videos of educational tutorials in other languages by generating lips from audio
- Full paper accepted at ICASSP 2019 [2], short paper published at CVPR Workshop 2018 (FIVER) [3]

PROJECT: **Assessor for Lipreader**
- Built a visual speech recognizer (lipreader) to classify spoken words
- Built an assessor for self-training on unlabelled data, zero-shot learning on out-of-vocabulary words

Other Experience

Scientist in Residence — *NextAI (startup accelerator)* — *April 2019 - present*
- Consultant for multiple startups on computer vision, deep learning and AI

Consultant, Computer Vision — *Playment (computer vision startup)* — *January 2018 - June 2018*
- Worked on semantic segmentation models for autonomous driving

Mentor — *Foundations of Artificial Intelligence and Machine Learning* — *January 2018 - May 2018*
IIIT Hyderabad, India — *six months certificate program for software professionals*
- Designed tutorials on machine learning, and mentored industry professionals.

Work Experience

Image Processing Engineer — *Embedded Systems Team* — *February 2016 - May 2017*
GreyOrange Robotics, India — *a multinational firm that designs, manufactures and deploys advanced robotics systems for automation at warehouses, distribution and fulfillment centres*

- Developed computer vision module to perform video processing in real time for warehouse automation
- Responsible for development and testing of entire code, including video processing module, module for communication with camera drivers, other systems, and server
- Research paper [4] based on work is published by ACM at ICIDE 2017

Associate Engineer — *Avionics Software & Systems Testing Group* *July 2014 - February 2016*
AIRBUS, INDIA — *a commercial aircraft manufacturer, the largest aeronautics & space company in Europe*

- 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)

THESIS PROJECTS

"De-fencing of Images using RGB-D Data" — M.Tech. Thesis *2013 - 2014*
IIT KHARAGPUR — *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 [5] based on work is published in the proceedings of ICAPR 2015
- Links — GitHub repository containing thesis, presentation, code files, and results

"Identification of Bilabial Consonants in Audio and Lip Closures in Video" — B.Tech. Thesis
IIT KHARAGPUR — *Prof. Rajiv Sahay, Department of Electrical Engineering* *2012 - 2013*

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

TALKS & OTHER ACHIEVEMENTS

- *May 2019* — Talk: Tutorial on GANs at the AI for Social Good Summer Lab, Montreal
- *Oct 2018* — Talk: "BigGAN - Large Scale GAN Training for High Fidelity Natural Image Synthesis" [presentation] — at Mila, University of Montreal, Canada
- *Feb 2018* — Talk: "Image de-fencing using RGB-D data" [presentation] — at Max Planck Insitute for Informatics, Saarbrücken, Germany
- *Feb 2018* — Talk: "Intuition behind LSTMs" [presentation] — at IIIT Hyderabad, India
- *Aug 2017* — Talk: "Mathematics of back-propagation in multi-layer perceptrons" [link] — at GreyOrange Robotics, India, and at IIIT-Hyderabad
- 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)

RESEARCH INTERNSHIPS

"Implementation of Carry-Free Arithmetic Operations in FPGA" *Summer 2013*
KU LEUVEN, BELGIUM — *Prof. Ingrid Verbauwhede, Computer Security & Industrial Applications*

- Designed arithmetic operations using Carry-Free Logic, simulated circuits in Xilinx
- Links — GitHub repository containing report, presentation, and related files

"Fingertip Gesture Recognizer using HMMs" *Summer 2012*
IIT KHARAGPUR, INDIA — *Prof. Aurobinda Routray, Department of Electrical Engineering*

- Implemented Hidden Markov Models (HMMs) in MATLAB, used to recognize shapes drawn by fingertip
- Links — GitHub repository containing report, presentation, code files, and results

"Measurement of Intra-die Power Variation in Sub-nm FPGA's" *Summer 2011*
IMPERIAL COLLEGE, LONDON — *Prof. Peter Cheung, Head, Electrical and Electronics Engineering*

- Measured the relative power consumption among the LookUp Tables (LUTs) of an FPGA
- Links — GitHub repository containing presentation, certificate, and recommendation letter

TECHNICAL SKILLS

Languages : C, C++, HTML/CSS, Javascript, Python, MATLAB, Shell, Verilog
Libraries: CUDA, IDS (cameras), Keras, LaTeX, OpenCV, PyTorch, Tensorflow
Operating Systems: OS X, Ubuntu, Windows