Avi Singh

PhD Student

750 Sutardja Dai Hall
Berkeley, CA 94720
✉ *avisingh@cs.berkeley.edu*
🖥 *http://people.eecs.berkeley.edu/ avisingh/*

Research Interests

Machine Learning, Robotics, Computer Vision

Education

2016 – **PhD in Computer Science** *University of California* Berkeley.
Advisor: Prof. Sergey Levine

2012–2016 **Bachelor of Technology** *Indian Institute of Technology* Kanpur.
Major: Electrical Engineering

Publications

Generalizing Sensorimotor Policies with Weakly Labeled Data
Avi Singh, Larry Yang, Sergey Levine.
ICCV 2017 (to appear). Also accepted at ICML 2017 Workshop on Lifelong Learning

arXiv link **Visual Dialog**
Abhishek Das, Satwik Kottur, Khushi Gupta, **Avi Singh**, Deshraj Yadav, Jose M. F. Moura, Devi Parikh, Dhruv Batra.
CVPR 2017 (Spotlight)

arXiv link **Recurrent Neural Networks for Driver Activity Anticipation via Sensory-Fusion Architecture**
Ashesh Jain, **Avi Singh**, Hema Koppula, Shane Soh, Ashutosh Saxena.
ICRA 2016

arXiv link **Brain4Cars: Car That Knows Before You Do via Sensory-Fusion Deep Learning Architecture**
Ashesh Jain, Hema S Koppula, Shane Soh, Bharad Raghavan, **Avi Singh**, Ashutosh Saxena.
IJRR 2016

Awards

2016 **EECS Department Fellowship**, UC Berkeley
2014-2015 **Academic Excellence Award**, IIT Kanpur (Awarded to 60 students out of 840).
2012-2013 **Academic Excellence Award**, IIT Kanpur (Awarded to 60 students out of 840).

Research Experience

August 2016 -

Graduate Student Researcher, *UC BERKELEY*
Deep Learning for Robotics and Computer Vision
under Prof. Sergey Levine.

Working at the intersection of machine learning, robotics, and computer Vision.
○ Building algorithms that allow robots to learn generalizable skills.

May-July 2016
project page

Research Intern, *VIRGINIA TECH*
Visual Dialog
under Prof. Dhruv Batra and Prof. Devi Parikh.

We are building AI systems that can have natural language conversations about images (in the form of sequential question answering). The work was presented at CVPR 2017. My contributions to the project are listed below:
○ Established the task as a retrieval-based problem for automated evaluation.
○ Developed an automated approach for collecting hard negatives for the retrieval problem.
○ Trained a *Sequence to Sequence* model based on LSTMs to generatively model the problem.
○ Trained retrieval models for the same problem, also based on LSTMs.

May-July 2015
project page

Research Intern, *CORNELL UNIVERSITY*
Brain4Cars: Anticipating Maneuvers via Learning Temporal Driving Models
under Prof. Ashutosh Saxena.

Brain4Cars addresses the problem of anticipating driver maneuvers several seconds before they happen. It fuses the information from driver-facing and road-facing cameras with data from other sensors to make its predictions. These predictions can then be passed to driver assistance systems that can warn the driver if the maneuver is deemed to be dangerous. My contributions to the project are listed below:
○ The KLT face tracker in the project was replaced with a facial landmark localization pipeline based on Constrained Local Neural Fields. This provided robust tracking and allowed the computation of head pose, which then served as a strong feature for maneuver anticipation.
○ Implemented Gaussian Mixture Model-based initialization, and LBFGS optimization for training Autoregressive Input Output Hidden Markov Models (AIOHMM), which are a modification of HMMs and used for anticipation in Brain4Cars.
○ The performance of the AIOHMM-based anticipation system improved from a Precision/Recall of 77.4/71.2 to 86.7/78.2.
○ Further testing on a Long Short Term Memory (LSTM) network (replacing the AIOHMM) increased the performance to a Precision/Recall of 90.5/87.4.

July-Dec 2014
Github-Stereo
Github-Mono

Undergraduate Researcher, *IIT-KANPUR*
Visual Odometry for Ground Vehicles
under Prof. KS Venkatesh.

Visual Odometry is the problem of estimating the trajectory and pose of a vehicle using a video stream from a camera (or a stereo pair) that is rigidly attached to the vehicle. Two **stereo** approaches were implemented and evaluated on the KITTI odometry benchmark:
○ Jiang2014: (Model Based ICP) 3D points triangulated from stereo data, inliers detected via the use of a Iterative Closest Point Algorithm that used a 1-DOF motion model for initial estimate. Efficient PnP algorithm is then applied on the selected inliers to obtain the final rotation and translation.
○ Howard2008: Inlier detection using an assumption of scene rigidity. Problem reduced to finding the maximum clique in a graph, solved using a heuristic. Levenberg-Marquardt used for minimizing the reprojection error on the selected inliers.

A **monocular** visual odometry approach was also implemented:
○ Utilizes Nister's five point algorithm for essential matrix computation.

Mentoring

Undergrads	Larry Yang (UC Berkeley), Dibya Ghosh (UC Berkeley)

Selected Hardware/Software Projects

October 2015
Github

Deep Learning for Visual Question Answering *For CS671*
under Prof. Amitabh Mukherjee, Department of Computer Science, IIT-Kanpur.
I developed Neural Network-based models to tackle the problem of answering open-ended natural language questions about images. I was able to achieve an accuracy of 53.34% on the test-dev split of the large-scale VQA dataset. This project was judged as the **best project** (out of 20 projects) in CS671 (NLP) course. I released the code for my approach, and it was the first open source project on the problem of VQA.

March 2014
Github

Hilbert Transform on FPGA/Verilog *For Techkriti 2014.*
Hilbert Transform is a mathematical operation used in Signal Processing. A hardware implementation can provide faster computations as compared to a generic CPU implementation. I Implemented a 32-point Discrete Hilbert Transform in **Verilog**, making use of the **Fast Fourier Transform** in the intermediate stages.

2012-14

Hardware Hacks at Electronics Club *Just-for-fun projects.*
- **Cashless Campus**: Developed an **arduino-based** point-of-sale device, with **biometeric (fingerprint) authentication** . An Arduino Mega was interfaced with an ethernet shield (with Wiznet51000 chip for UDC/TCP stacks), an LCD, a touchscreen, a thermal printer, and an SD card via **SPI and UART**.[Github]
- **SNAKE64**: Implemented the classic 'Snakes' game on a self-fabricated LED matrix of size 8x8, and wrote an original C code for ATmega8 to drive the matrix.
- **Laser Tag**: An infrared-based gun was implemented using IR LEDs and 38KHz modulated wave was generated on an ATmega32. The receiving unit employed a TSOP to detect bursts of infrared sent from the gun.

Press Coverage

MIT Tech Review	This Car Knows Your Next Misstep Before You Make It link
MIT Tech Review	How Future Cars Will Predict Your Driving Maneuvers Before You Make Them link
Fortune	Better driving through artificial intelligence link

Relevant Coursework at IIT-Kanpur

Vision/ML/AI	Machine Learning for Computer Vision, Learning with Kernels, Natural Language Processing, Online Learning and Optimization, Probabilistic Mobile Robotics, Convex Optimization, Artificial Intelligence, Applied Game Theory
Mathematics	Linear Algebra, Probability and Statistics, Multivariate Calculus, Ordinary/Partial Differential Equations, Complex Analysis
Algorithms	Algorithms-II, Data Structures and Algorithms, Intro to Computing
Signal Processing	Signals and Systems, Digital Signal Processing, Communication/Information Theory, Principles of Communication

Relevant Coursework at UC Berkeley

Vision/ML/AI Deep Reinforcement Learning, Advanced Deep Learning Seminar, Vision Reading Group, Statistical Learning Theory, Optimization Algorithms and Analysis, Algorithmic Human Robot Interaction

Technical Skills

Languages C, C++, Python, Lua, MATLAB

Libraries Tensorflow, Torch, Keras, OpenCV, ROS

OS GNU/Linux (Ubuntu), Microsoft Windows

Other Git, LaTeX

Activities at IIT-Kanpur

2014-15 **Electronics Club Coordinator, Science and Technology Council, IIT-Kanpur**.
- Floated, mentored and ensured the completion of nine summer projects including a **3D Laser Scanner**, **A Video Surveillance Robot**, Conway's Game of Life simulation using FPGAs, Fast Fourier Transform on FPGA, An accelerometer based fitness and sleep tracker with accompanying Android App, a Surveillance system with **face recognition**, and a Laser Tag system.
- Led a team of 16 secretaries and handled a budget of Rs.76,000 to organize lectures, workshops, competitions, and another Rs.474,000 for funding projects and for participation in external events.
- Lectures attended by 400+ people, workshops attended by 200+ people, and participation of 100+ people in Takneek (intra-IIT Kanpur technical festival) Electronics competitions.

2014- **Blogger**.
I love explaining things in a simple manner, and I've written several blog posts that introduce research problems (and ways to tackle them) to beginners. Some of my posts have been published on websites like **KDNuggets** and **LearnOpenCV**. Some recent posts:
- Deep Learning for Visual Question Answering link
- Monocular Visual Odometry in OpenCV link
- Visual Odometry from Scratch - A tutorial for beginners link

2015-16 **Member, Special Interest Group in Machine Learning (SIGML), IIT-Kanpur**.
I have given a talk about my work, and attended lectures given by others in the various sub-fields of Machine Learning.

2013-14 **Coordinator, ECDC, Techkriti 2014**.
ECDC is the name given to Electronics Circuit Design Competitions held at Techkriti 2014, IIT-Kanpur's annual inter-college tech festival.
- Designed and verified the problem statement for the competition Electromania, which was about building an Infrared-based communication system for short distances.
- Prepared sample codes and tutorials for the participants.

2013-14 **Secretary, Electronics Club**.
Assisted in organization of lectures, workshops, tutorials, and maintenance of club.

2013-16 **Student Guide, Counseling Service**.
- Helped six freshmen in settling in the new college environment.
- Assisted in the organization of various Counseling Service activities such as the Orientation Program.