

RESEARCH INTERESTS

Computer Vision, Machine Learning, Robotics

EDUCATION

Indian Institute of Technology
Kanpur

GPA: **9.4/10** (after six semesters) • Bachelor of Technology (expected July 2016)
Major: Electrical Engineering • Minor: Computer Science (Artificial Intelligence)

PUBLICATIONS

[Recurrent Neural Networks for Driver Activity Anticipation via Sensory-Fusion Architecture](#)

<http://arxiv.org/abs/1509.05016>

Ashesh Jain, **Avi Singh**, Hema Koppula, Shane Soh, Ashutosh Saxena

under review at IEEE International Conference on Robotics and Automation (ICRA), Stockholm, Sweden, 2016

[Brain4Cars: Sensory-Fusion Recurrent Neural Networks for Driver Activity Anticipation](#)

[BayLearn.org](#)

Ashesh Jain, Shane Soh, Bharad Raghvan, **Avi Singh**, Hema Koppula, Ashutosh Saxena

Accepted as **Full ORAL** at Baylearn Symposium, Menlo Park, California, USA, October 2015

RESEARCH INTERNSHIP AT CORNELL UNIVERSITY

[Brain4Cars - Anticipating Maneuvers via Learning Temporal Driving Models](#)

MAY 2015 - JULY 2015

under Prof. Ashutosh Saxena, Robot Learning Lab, Department of Computer Science, Cornell University

www.brain4cars.com

Brain4Cars takes input from driver-facing and road-facing cameras, and predicts what maneuvers the driver is going to perform in the next few seconds. My contribution to the project was in the following areas:

- Revamped the vision pipeline of the project by incorporating a Constrained Local Neural Field model for enhanced head pose tracking of the driver.
- Implemented Gaussian Mixture Model-based initialization, and LBFGS optimization for training Autoregressive Input Output Hidden Markov Models.
- The above changes improved the performance of the AIOHMM-based anticipation system from a Precision/Recall of 77.4/71.2 to 86.7/78.2.
- Further testing on a Recurrent Neural Network with Long Short Term Memory (LSTM) cells increased the performance to a Precision/Recall of 90.5/87.4.

RESEARCH EXPERIENCE AT COMPUTER VISION LAB, EE, IIT-KANPUR

[Visual Odometry for Ground Vehicles](#)

AUG 2014 - DEC 2014

under Prof. KS Venkatesh, Computer Vision Lab, Department of Electrical Engineering, IIT-Kanpur

Two **stereo** approaches were implemented and evaluated on KITTI odometry benchmark:

- Jiang2014: Model Based ICP: 3D points triangulated from stereo data, inliers detected via the use Iterative Closest Point Algorithm that uses a 1-DOF motion model for initial estimate. Efficient PnP Algorithm is then used 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. Implementation available on [Github](#).

A **monocular** visual odometry approach was also implemented:

- Utilizes Nister's five point algorithm for essential matrix computation. Implementation available on [Github](#).

[Estimation of Scene Flow from RGBD data](#)

MAY 2014 - JULY 2014

under Prof. KS Venkatesh, Computer Vision Lab, Department of Electrical Engineering, IIT-Kanpur

- Designed and implemented two approaches for estimation of Scene Flow from RGB-D data, one using 'Total Least Squares Solution', the other utilizing 'Global Minimum Energy Solution'.
- Captured RGBD Data from a Microsoft Kinect using OpenNI and OpenCV libraries in C++, and qualitatively evaluated the results.

SELECTED PROJECTS

[Deep Learning for Visual Question Answering](#)

SEPTEMBER 2015 - OCTOBER 2015

COURSE PROJECT FOR CS671, under Prof. Amitabha Mukherjee, Department of Computer Science, IIT-Kanpur

[Blog Post](#)

I developed Neural Network-based models to tackle the problem of answering open-ended questions about images.

- The image is passed through a pre-trained Convolutional Neural Network, and the activation from the layer just before the softmax layer are extracted.
- Every word is converted to its word vectors using the Glove Word Embeddings, and these embeddings are sequentially passed to a Long Short Term Memory Network (LSTM). The output of the LSTM is taken as the an embedding for the question.
- The image embedding and the question embedding are concatenated and passed through two fully connected layers with 50% dropout.
- The VQA dataset was used with 248K questions for training and 121K questions for validation. The system was able to achieve an accuracy of 47.8% on the validation set when the experimnt was run for 100 epochs without any hyperparameter tuning.
- Implementation has been open-sourced on [Github](#) under the MIT License, is the first open-source project to work with the VQA dataset and has 70 stars.

[DAAnT - Dental Automated Analysis Tool \(Computer Vision for intraoral cameras\)](#)

JANUARY 2015

MIT REDX CAMP, under Dr. Hyunsung Park, Postdoc at Camera Culture, MIT Media Lab

We developed Computer Vision algorithms for diagnosis of dental diseased from images obtained from an intraoral camera. My contribution in the eight-member team was primarily in the following areas:

- **Stitching Images:** Used Affine SIFT to overcome the difficulties in feature matching due to changes in perspective, and then computed Homography with RANSAC to stitch the images.
- **Segmenting Every tooth:** Marker-controlled watershed transform was used to segment every tooth from the image. Both semi-automated and automated approaches were implemented and compared.

FOR CS679, UNDER PROF. VINAY NAMBOODIRI, DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING, IIT-KANPUR

- Hidden Conditional Random Fields were used for human activity classification from RGBD sequences.
- Pose features were extracted from a skeleton detection pipeline
- A novel normalization operation was developed for reducing the effect of variance in body sizes of different subjects
- An accuracy of 71% was achieved on a reduced version of MSR Daily Activity 3D Dataset (6-class classification)

Landmark-based Robotic Localisation from RGBD data

DECEMBER 2013

UNDER MR. ARJUN BHASIN, PROJECT ENGINEER AT MECHATRONICS LAB, IIT-KANPUR

- Geometric Triangulation was used to determine the pose (modelled as a set of random variables with Gaussian Distribution) of a robot, from information obtained from noisy landmarks.
- A Microsoft **Kinect** was used to identify the landmarks (using color histogram based models), **CAMshift** algorithm was used to track these landmarks, and bearing measurements were calculated using the depth data.
- An error model for the data obtained from a Kinect was used along with the Error Propagation Law to arrive at the uncertainty in the final pose computed using the Geometric Localisation Algorithm.

Hilbert Transform on FPGA/Verilog

MARCH 2014

For Techkriti 2014

[\[Github\]](#)

- Implemented a 32-point Discrete Hilbert Transform in **Verilog**, making use of the **Fast Fourier Transform** in the intermediate stages.
- Won **2nd position** in FPGA Design, Techkriti 2014.

Hardware Hacks
@ Electronics Club

Cashless Campus: Developed an **arduino-based** point-of-sale device, with **biometric (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.

SCHOLASTIC ACHIEVEMENTS AND AWARDS

2014 Awarded an **A* grade**, for exceptional performance in **Undergraduate Project - 1**

2012-2013 **Academic Excellence Award**, IIT Kanpur (Awarded to 60 students out of 840).

MAY 2012 Secured All India Rank **387** in IIT-JEE 2012 out of **0.5m** candidates

MAY 2012 Secured All India Rank **345** in AIEEE 2012 out of **1.2m** candidates

JAN 2012 **Top 1%** in National Physics Olympiad

RELEVANT COURSES COMPLETED AND ONGOING(*)

Machine Learning, AI: Machine Learning for Computer Vision, Learning with Kernels*, Natural Language Processing*, Applied Game Theory

Mathematics: Linear Algebra, Probability and Statistics, Multivariate Calculus, Ordinary/Partial Differential Equations, Complex Analysis

Other: Algorithms-II*, Data Structures and Algorithms, Digital Signal Processing, Communication/Information Theory

TECHNICAL SKILLS

Programming Languages	C, C++, MATLAB, Python	Microcontrollers	Atmel ATmega, dsPIC, Arduino
Software Packages	OpenCV, ROS, Point Cloud Library (PCL)	Imaging Systems	Microsoft Kinect, PST-110
Operating Systems	GNU/Linux (Ubuntu), Microsoft Windows	Other Hardware	Raspberry Pi, Beagleboard, Servos, IMUs
Other	Git, \LaTeX		

ACTIVITIES

Electronics Club Coordinator, Science and Technology Council, IIT-Kanpur

2014-2015

- 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.
- Leading a team of **16 secretaries** and handling a budget of **Rs.76,000** to organize lectures, workshops, competitions, and another **Rs.4,74,000** 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.

Member, Special Interest Group in Machine Learning (SIGML), IIT-Kanpur

2015 - Present

I have given a talk about my work, and attended lectures given by others in the various sub-fields of Machine Learning.

Coordinator, ECDC, Techkriti 2014

2013 - 2014

Designed and verified the problem statement for the event Electromania. Prepared sample codes and tutorials for the participants.

Secretary, Electronics Club

2013 - 2014

Assisted in organisation of lectures, workshops, tutorials, and maintenance of club.

Student Guide, Counselling Service

2013 - Present

Assisted in the organisation of various counselling service activities such as the Orientation Program and helped a group of six freshers in settling in the new college environment.