

# Isht Dwivedi

718-679-3159 | [isht.dwivedi@gmail.com](mailto:isht.dwivedi@gmail.com) | [Google Scholar](https://scholar.google.com/citations?user=IshtDwivedi) | [github.com/isht7](https://github.com/isht7)

## EDUCATION

---

### Columbia University

*M.S. in Computer Science, Specialization: Machine Learning, GPA: 3.94/4.00*

New York, USA

Aug 2017 – Dec 2018

### Indian Institute of Technology Guwahati

*B. Tech., Major: Electronics and Electrical Engg., Minor: Computer Science, GPA: 8.36/10*

Guwahati, India

Jul 2012 – Jun 2016

## WORK EXPERIENCE

---

### Honda Research Institute USA

*Research Engineer*

Feb 2019 – Present

San Jose, USA

- Road scene understanding with emphasis on unstructured events like construction zones. Bird's Eye View (BEV) semantic segmentation map is directly predicted from perspective view video stream captured using a front-facing camera mounted on a car. Used COLMAP to create colored 3D point clouds from video snippets which were annotated to synthesize ground truth. Achieved state of the art performance on BEV segmentation. (Paper submitted to BMVC 2021).
- Future trajectory prediction for all agents in a scene in a single shot in constant time. Single shot nature of this work makes it faster than other works in crowded situations (published at **IROS 2020**). Single shot nature achieved by predicting composite fields from neural network. Non-Local interaction block used to capture interactions between agents. Semantic segmentation features used to make semantically aware predictions.
- Improved road place classification (see intern project below) performance by redesigning neural network architecture and adding semantic segmentation as auxiliary task. F-score improved from 28% to 40%.
- Implement road scene understanding algorithms (road place classification, construction zone detection) on a car to obtain predictions in real time.

### Honda Research Institute USA

*Research Intern, Dr. Behzad Dariush*

May 2018 – Aug 2018

Mountain View, USA

- Worked on road place and condition classification for videos obtained from a front facing camera mounted on a car. Place classification includes classification of road scenes, road surface conditions, road weather and road type (paper published in **ICRA 2019**). Two stage neural network used – first stage selects candidate video snippet from long video, second stage classifies the snippet.

### Active Interpretation of Disparate Alternatives (AIDA) Project

*Research credit with Prof. Shih-Fu Chang, Columbia University*

Jan – May 2018

New York City, USA

- Worked on the AIDA project with Prof. Shih-Fu Chang in the DVMM lab.
- My work involved discovering relevant visual concepts from a given scenario (body of text). Using the discovered visual concepts, I experimented with different methods to create a weakly supervised object detection method for these visual concepts using the Open Images dataset.

### Indian Institute of Science, Video Analytics Lab

*Research Intern, Prof. Venkatesh Babu*

Aug 2016 – May 2017

Bangalore, India

- Hierarchical CNN used for semantic segmentation of hand drawn sketches with pose estimation as auxiliary task.
- Fully automatic, works without class information on over 11 classes including cat, car, bird, bicycle and airplane.
- Sketches not required for training: Sketchification process used to convert natural photos to sketch like images.
- Published at **ACMMM 2017**.

### Ecole Polytechnique, Laboratoire Leprince-Ringuet

*Summer Intern, Prof. Vincent Boudry*

May 2015 – Jul 2015

Paris, France

- High energy particles decay in a particle detector generating new particles in a graph like pattern (shower). Designed and implemented an approach to reconstruct vertices of this graph.
- Written in C++, using PCA and Arbor Particle Flow Algorithm, experiments done on Monte Carlo simulations.

## PUBLICATIONS

---

Reza Ghoddoosian, **Isht Dwivedi**, Nakul Agarwal, Behzad Dariush. “*Weakly-Supervised Online Action Segmentation in Multi-View Instructional Videos*” **accepted** at the Conference on Computer Vision and Pattern Recognition (**CVPR**) **2022**.

**Isht Dwivedi**, Srikanth Malla, Yi-Ting Chen, Behzad Dariush. “*Bird Eye View Segmentation Using Lifted 2D Semantic Features*”, British Machine Vision Conference **published at (BMVC) 2021**. [Link to paper](#).

**Isht Dwivedi**, Srikanth Malla, Behzad Dariush, Chiho Choi. “*SSP: Single Shot Future Trajectory Prediction*”, International Conference on Intelligent Robots and Systems (**IROS**) **2020**. [Link to paper](#).

**Isht Dwivedi\***, Athmanarayanan Narayanan\*, Behzad Dariush. “*Dynamic Traffic Scene Classification with Space-Time Coherence*”, International Conference on Robotics and Automation (**ICRA**) **2019**. [Link to paper](#).

Kaveri A Thakoor, Qian Zheng, Linyong Nan, Xinhui Li, Emmanouil Manos Tsamis, Rashmi Rajshekhar, **Isht Dwivedi**, Iddo Drori, Paul Sajda, Donald C Hood. “*Assessing the Ability of Convolutional Neural Networks to Detect Glaucoma from OCT Probability Maps*”, Journal on Investigative Ophthalmology & Visual Science (**IOVS**) **2019**. [Link to paper](#).

Iddo Drori, **Isht Dwivedi**, Pranav Shrestha, Jeffrey Wan, Yueqi Wang, Yunchu He, Anthony Mazza, Hugh Krogh-Freeman, Dimitri Leggas, Kendal Sandridge, Linyong Nan, Kaveri Thakoor, Chinmay Joshi, Sonam Goenka, Chen Keasar, Itsik Pe’er. “*High-Quality Protein Q8 Secondary Structure Prediction by Diverse Neural Network Architectures*”, **NeurIPS 2018 Workshop** on Machine Learning for Molecules and Materials. [Link to paper](#).

Ravi Kiran, **Isht Dwivedi**, Abhijat Biswas, Sahil Manocha, V. Babu. “*SketchParse: Towards Rich Descriptions for Poorly Drawn Sketches Using Multi-Task Deep Networks*”, ACM Multimedia (**ACM MM**) **2017**. [Link to paper](#).

**Isht Dwivedi**, Swapnil Gupta, V. Venugopal, S. Sundaram. “*Online Writer Identification using Sparse Coding and Histogram based Descriptors*”, **Oral** presentation at International Conference on Frontiers in Handwriting Recognition (**ICFHR, oral**) **2016**. [Link to paper](#).

Srikanth Malla, **Isht Dwivedi**, Behzad Dariush, Chiho Choi. “*NEMO: Future Object Localization Using Noisy Ego Priors*”. [Link to paper](#).

## REVIEWER FOR JOURNALS & CONFERENCES

---

Reviewer for the Conference on Computer Vision and Pattern Recognition (**CVPR**) **2022**.

Reviewer for the European Conference on Computer Vision (**ECCV**) **2022**.

Reviewer for the IEEE Internet of Things Journal (**IoT**) **2022**.

Reviewer for the International Conference on Intelligent Robots and Systems (**IROS**) **2022**.

Reviewer for The IEEE Journal on Robotics and Automation Letters (**RA-L**) in **2022**.

Reviewer for The Workshop on Autonomous Driving at CVPR **2022**.

Reviewer for The Workshop on Robustness in Sequential Data at CVPR **2022**.

## PATENTS

---

US Patent 16374205, “Scene classification”, Athmanarayanan Lakshmi Narayanan, **Isht Dwivedi**, Behzad Dariush

US Patent 11034357, “Scene classification prediction”, Athmanarayanan Lakshmi Narayanan, **Isht Dwivedi**, Behzad Dariush

US Patent 16917864, “Composite field based single shot prediction”, **Isht Dwivedi**, Chiho Choi, Srikanth Malla, Behzad Dariush

Submitted patent application to US Patent office for work related to Bird’s Eye View segmentation (2021).

Submitted patent application to US Patent office for work related to Temporal Action segmentation in videos (2021).

## RESEARCH PROJECTS

---

### **Automatic Image Colorization**

Oct – Dec 2017

*Deep Learning Course Project*

*New York City, USA*

- Automatic Image Colorization of greyscale images using a CNN trained on Places365 dataset.
- Modified MSE loss using attention in images to improve colorization, classification used as an auxiliary task.

### **Online writer identification**

Sept – May 2016

*Bachelor Thesis Project*

*Guwahati, India*

- Improvement of 10% obtained over state-of-the-art results for the IAM On-Line Handwriting Database when using only online features at the text-line level.
- Developed novel histogram-based features to represent handwriting characteristics of a writer.
- Codebook developed using sparse coding. Modified Tf-Idf approach used to create document descriptor. One vs. all SVM model used as classifier. Paper (oral) published at **ICFHR 2016**.

### **Snake Robot with passive wheels and active joints** [[Video of robot](#)]

Jan – Apr 2015

*Design Course Project*

*Guwahati, India*

- Designed and developed a wirelessly controlled snake robot with passive wheels and active joint modules. Motion of robot specifically designed to imitate that of a snake.
- Zigbee protocol used for wireless communication and Arduino Mega 2560 used for control of servo motors.

## TEACHING EXPERIENCE

---

**Teaching Assistant, CS 4995: Deep Learning**, Columbia University, Spring 2018

*Instructor:* Prof. Sameer Maskey

**Teaching Assistant, CS 4995: Deep Learning**, Columbia University, Fall 2018

*Instructor:* Prof. Iddo Drori

## RELEVANT COURSEWORK

---

**Graduate level courses:** Deep Learning, Machine Learning, Natural Language Processing, Unsupervised Methods of Learning, Bayesian Models for Machine Learning, Visual Interfaces to Computers

**Undergraduate courses:** Computer Vision, Image Processing, Pattern Recognition and Machine Learning, Probability and Random Processes, Digital Signal Processing, Linear Algebra and Real Analysis, Differential Equations and Multi-variable Calculus, Speech Technology

## FELLOWSHIPS

---

CA Fellow in Spring 2019 semester which allowed for full tuition waiver.

## TECHNICAL SKILLS

---

**Programming Languages:** Python, C/C++, MATLAB

**Tools and Technologies:** Pytorch, Tensorflow, Colmap, Caffe