

Tanmay Gupta

Siebel Center, 201 N Goodwin Ave, Urbana, Illinois 61801, USA
tgupta6@illinois.edu • +1 (217) 419-0139 • <http://tanmaygupta.info>

RESEARCH INTERESTS

Learning representations for Vision and Language

Interested in building intelligent systems that combine visual knowledge with natural language understanding to enhance the way computers “see” our world.

EDUCATION

University of Illinois Urbana Champaign, Illinois, USA

Doctoral Candidate in Computer Science

Aug 2014 – Present

Adviser: Prof. Derek Hoiem

GPA: 3.87 / 4.00

Indian Institute of Technology Kanpur, Uttar Pradesh, India

Bachelor of Technology (B.Tech.) in Electrical Engineering

Jul 2010 – May 2014

CPI (Cumulative Performance Index): 9.9 / 10.0

PUBLICATIONS

T. Gupta, A. Schwing, D. Hoiem. “Seeing Human-Object Interactions Through Eyes Of Object And Pose Detectors”. **NIPS 2018 Submission**

T. Gupta, D. Schwenk, A. Farhadi, D. Hoiem, A. Kembhavi. “Imagine This! Scripts to Compositions to Videos”. **ECCV 2018**

T. Gupta, K.J. Shih, S. Singh, D. Hoiem. “Aligned Image-Word Representations Improve Inductive Transfer Across Vision-Language Tasks”. **ICCV 2017**

T. Gupta, D. Shin, N. Sivagnanasadan, D. Hoiem. “3DFS: Deformable Dense Depth Fusion and Segmentation for Object Reconstruction from a Handheld Camera”. **arXiv 2016**

J. Rock, T. Gupta, J. Thorsen, J. Gwak, D. Shin, D. Hoiem. “Completing 3D Object Shape from One Depth Image”. **CVPR 2015**

T. Gupta, S. Gupta, A.K. Jagannatham. “Face Tracking and Recognition with Orientation, Pose and Illumination Variations”. **UG Thesis 2014**

INTERNSHIPS

Research Intern, Allen Institute for Artificial Intelligence

May 2017 – Dec 2017

Collaborators: Aniruddha Kembhavi, Dustin Schwenk, Ali Farhadi, Derek Hoiem

- Proposed and implemented a semi-parametric approach for text to video generation. We demonstrated this approach for a new richly annotated FLINTSTONES dataset.

Software Development Intern, Visual Search Team, A9.com

May 2015 – Aug 2015

Manager: C. J. Taylor, *Mentor:* Michael Lou

- Developed an algorithm for depth estimation and foreground segmentation from a video stream.

Undergraduate Research Intern, Cornell University

May 2013 – Jul 2013

Supervisor: Prof. Tsuhan Chen, Professor and Director of ECE, Cornell University

- Studied and implemented algorithms for correspondence estimation and alignment of 3D point clouds.

RESEARCH

Vision and Language

2016 – Present

- Developed a system for human-object interaction detection that doubles the state-of-the-art on HICO-Det dataset while using pretrained object and pose detectors. **NIPS 2018 Submission**

- Proposed an iterative composition and retrieval approach for text to video generation. **ECCV 2018**

- Developed a shared vision-language representation that enhances inductive transfer across vision-language tasks like visual recognition and VQA. **ICCV 2017**

3D Reconstruction

2014 – 2016

- Developed a system for creating a 3D mesh of an object from an RGB video. Contributions include an improved cost function for dense depth estimation, robust implicit surface estimation with zero-crossing correction, and joint 2D-3D segmentation. **arXiv 2016**

- Developed a data-driven approach for 3D mesh reconstruction from a single depth image. Contributions include symmetry constrained thin-plate splines mesh deformation. **CVPR 2015**

Face Detection, Tracking and Recognition

2013 – 2014

- Designed an algorithm for pose and illumination invariant face detection, tracking and recognition. Delivered a final presentation at the prestigious Bhabha Atomic Research Center. **UG Thesis**

AWARDS	AI2 Award 2017 Allen Institute for Artificial Intelligence Received a \$10K award in support of my work on semi-parametric text to video generation
	Sridhar Memorial Prize 2014 IIT Kanpur Given to the best student of B.Tech final year Electrical Engineering at the end of 6th Semester
	Todai-IIT Scholarship (Awarded Twice) 2011 – 2013 University of Tokyo and Mori Seiki Company, Japan Given to 8 students each from 5 major IITs for academic excellence
	Certificate of Merit for Academic Excellence (Awarded Thrice) 2010 – 2013 IIT Kanpur, India Given to top 5% students in each department at IIT Kanpur for academic excellence
TALKS	Enhancing Inductive Transfer in Vision-Language Tasks May 2017 <i>Venue:</i> Midwest Vision Workshop, Here Research, Chicago
	Role of Language in Vision Apr 2017 <i>Venue:</i> CS543: Computer Vision Course, UIUC
	Tensorflow Tutorial Mar 2017 <i>Venue:</i> Virtusense Technologies
	Deep Learning Panel Nov 2016 <i>Venue:</i> Big Data Summit, Research Park, UIUC
EMPLOYMENT & OUTREACH	Graduate Research Assistant Aug 2014 – Present <i>Supervisor:</i> Prof. Derek Hoiem, Department of Computer Science, UIUC
	Organizer, Vision Lunch Aug 2016 – May 2017 Primary vision and deep learning reading/presentation group in Department of Computer Science, UIUC
	Graduate Teaching Assistant Jan 2017 – May 2017 Prepared and delivered a lecture on Vision-Language research, held discussion hours, graded homeworks, and answered questions on Piazza for graduate computer vision course CS543 offered by Prof. Derek Hoiem.
	Senior Student Research Associate Aug 2013 – Dec 2013 <i>Supervisor:</i> Prof. Aditya K. Jagannatham, Department of Electrical Engineering, IIT Kanpur
	Student Research Associate Nov 2012 – Apr 2013 <i>Supervisor:</i> Prof. Aditya K. Jagannatham, Department of Electrical Engineering, IIT Kanpur
	Assistant Coordinator, Junkyard Wars Feb 2011 Competition in Techkriti, the Annual Inter College Technical Festival at IIT Kanpur
PROJECTS	Personal Blog [https://bigredt.github.io] Dec 2013 – Present <ul style="list-style-type: none"> • Technical blog about AI, Computer Vision and related topics.
	pyAIUtils [https://github.com/aiUIUC/pyAIUtils] Apr 2016 – Dec 2016 <ul style="list-style-type: none"> • Primary contributor to this code repository for utility functions and wrappers built around Tensorflow
	Efficient End-to-End Training of CRFs using Incomplete Data Dec 2016 <ul style="list-style-type: none"> • Proposed an efficient gradient descent algorithm for learning CRFs using Expectation Maximization and Naive Mean Field approximation. • Applied the algorithm for learning a CRF for image segmentation using a trimap with unlabelled regions.
	Stochastic Gradient Descent with Importance Sampling April 2016 <ul style="list-style-type: none"> • Studied recently proposed (Needell et.al NIPS 2014) convergence guarantees for optimization of functions with certain properties using SGD with importance sampling to create mini-batches. • Applied the method to the novel case of L2-normalized logistic regression and demonstrated significantly superior convergence performance with respect to SGD.

Semantic Role Labeling using Recurrent Neural Networks

Dec 2015

- Trained and evaluated different RNN architectures using Torch for a simplified SRL task of predicting the argument structure corresponding to a given predicate in a sentence.
- Experimentally evaluated the effect of activations (Sigmoid and ReLU), number of layers and hidden units for simple RNN and LSTM architectures using GloVe word representations.

Adaptive Crowd-sourcing via EM with Prior

Apr 2015

- Derived the Expectation-Maximization update rules for shifted and rescaled Beta prior
- Developed an algorithm for adaptively assigning tasks to workers for more efficient use of an edge budget

Implementation of Subdivision Surface

Dec 2014

- Implemented Catamull-Clark subdivision surface algorithm in VC++ using half-edge data structure
- Rendered the object using Phong shading model and texture mapping

TECHNICAL SKILLS

Prog. Languages: Python, MATLAB, C++, Java

Libraries: Pytorch, TensorFlow, OpenCV, OpenGL, Weka

Development Environments: VS Code, Emacs, Visual Studio, NetBeans, Git

RELEVANT COURSES

Machine Learning: Inference in Graphical Models, Cutting-Edge Trends in Deep Learning and Recognition, Computational Inference and Learning (P. Moulin), ML (D. Roth), ML for Signal Processing (P. Smaragdis), ML Tools and Techniques (H. Karnick), Mathematics for ML (H. Karnick)

Computer Vision and Graphics: Computer Vision, Interactive Computer Graphics, Image Processing, Digital Signal Processing

CS Theory: Algorithms (Grad), Data Structures and Algorithms

Mathematics: Matrix Theory and Linear Estimation, Probability and Statistics, Statistical Simulation and Data Analysis, Convex Optimization, Nonlinear Optimization