

## EDUCATION

- **University of California, Davis** (Fall, 2015 - Current)
  - PhD in Computer Science Advisor: [Prof. Yong Jae Lee](#)
- **Robotics Institute, Carnegie Mellon University, USA** (August 2013 – December 2014)
  - Masters in Robotics QPA: **4.05** Advisors: [Prof. Alexei Efros](#), [Prof. Kayvon Fatahalian](#)
- **International Institute of Information Technology (IIIT), Hyderabad, India** (August 2009 – May 2013)
  - B.Tech ( Honours ) in Computer Science and Engineering GPA: **9.07/10** Advisor: [Prof. P. J. Narayanan](#)

## INTERESTS

- **Computer Vision and Machine Learning** *esp.* to develop large scale data driven methods for visual recognition systems with minimal human supervision.

## PUBLICATIONS

- **FineGAN: Unsupervised Hierarchical Disentanglement for Fine-Grained Object Generation and Discovery:** Krishna Kumar Singh\*, Utkarsh Ojha\*, Yong Jae Lee. [Arxiv](#). [\[Project Page\]](#).
- **Hide-and-Seek: A Data Augmentation Technique for Weakly-Supervised Localization and Beyond:** Krishna Kumar Singh, Hao Yu, Aron Sarmasi, Gautam Pradeep, Yong Jae Lee. [Arxiv](#). [\[Paper\]](#).
- **You reap what you sow: Generating High Precision Object Proposals for Weakly-supervised Object Detection:** Krishna Kumar Singh, Yong Jae Lee. [\[Under submission\]](#)
- **DOCK: Detecting Objects by transferring Common-sense Knowledge:** Krishna Kumar Singh, Santosh Divvala, Ali Farhadi, Yong Jae Lee. *Appeared in* [European Conference on Computer Vision \(ECCV\), 2018](#). [\[Project Page\]](#)
- **Who Will Share My Image? Predicting the Content Diffusion Path in Online Social Networks:** Wenjian Hu, Krishna Kumar Singh\*, Fanyi Xiao\*, Jinyoung Han, Chen-Nee Chuah, Yong Jae Lee. *Appeared in* [ACM International Conference on Web Search and Data Mining \(WSDM\), 2018](#). [\[Paper\]](#) (\* equal contribution)
- **Hide-and-Seek: Forcing a Network to be Meticulous for Weakly-supervised Object and Action Localization:** Krishna Kumar Singh, Yong Jae Lee. *Appeared in* [International Conference on Computer Vision \(ICCV\), 2017](#). [\[Project Page\]](#)
- **Identifying First-Person Camera Wearers in Third-Person Videos:** Chenyou Fan, Jangwon Lee, Mingze Xu, Krishna Kumar Singh, Yong Jae Lee, David J. Crandall, Michael S. Ryoo. *Appeared in* [IEEE Conference on Computer Vision and Pattern Recognition \(CVPR\), 2017](#). [\[Paper\]](#)
- **End-to-End Localization and Ranking for Relative Attributes:** Krishna Kumar Singh, Yong Jae Lee. *Presented in* [European Conference on Computer Vision \(ECCV\), 2016](#). [\[Project Page\]](#)
- **Track and Transfer: Watching Videos to Simulate Strong Human Supervision for Weakly-Supervised Object Detection:** Krishna Kumar Singh, Fanyi Xiao, Yong Jae Lee. *Presented in* [IEEE Conference on Computer Vision and Pattern Recognition \(CVPR\), 2016](#). [\[Project Page\]](#)
- **KrishnaCam: Using a Longitudinal, Single-Person, Egocentric Dataset for Scene Understanding Tasks:** Krishna Kumar Singh, Kayvon Fatahalian, Alexei A. Efros. *Presented in* [Winter Conference on Applications of Computer Vision \(WACV\), 2016](#). [\[Project Page\]](#)
- **Storytelling Patches: Predicting Tourist Spots in a City:** Aayush Bansal, Krishna Kumar Singh. *Presented in* ECCV 2014, Workshop on Storytelling with Images and Videos (VisStory). [\[Paper\]](#)
- **Geometry Directed Browser for Personal Photographs:** Aditya Deshpande, Siddharth Choudhary, P J Narayanan, Krishna Kumar Singh, Kaustav Kundu, Aditya Singh and Apurva Kumar. *Appeared in* [Eighth Indian Conference On Vision, Graphics And Image Processing \(ICVGIP\), 2012 \(Oral\)](#). [\[Paper\]](#)
- **Hybrid Multi-Core Algorithms for Regular Image Filtering Applications:** Shrenik Lad, Krishna Kumar Singh, Kishore Kothapalli and P.J. Narayanan. *Appeared in* [International Conference on High Performance Computing \(HiPC\) Student Research Symposium, 2012](#). [\[Paper\]](#)

## WORK AND TEACHING EXPERIENCE

- Research intern, Allen Institute for Artificial Intelligence (AI2) (Summer, 2017) Advisors: [Dr. Santosh Divvala](#), [Prof. Ali Farhadi](#). Worked on improving object detection by transferring common-sense knowledge.
- Computer Vision intern at Intel Labs (Summer, 2015): Worked on video summarization and retrieval.
- Graduate Research Assistant at UC, Davis (Fall 2015 – Current) Advisor: [Prof. Yong Jae Lee](#).
- Graduate Research Assistant at RI, CMU (Spring 2014 - 2015) Advisors: [Prof. Alexei Efros](#), [Prof. Kayvon Fatahalian](#).
- Research Scholar at RI, CMU (Summer, 2012) Advisor: [Prof. Martial Hebert](#).
- Research Assistant at IIIT-Hyderabad (Summer 2011) and developed DLD virtual lab. Advisor: [Prof. P. J. Narayanan](#).
- Teaching Assistant at IIIT-Hyderabad for Information Retrieval and Extraction, Data Structures and Cloud Computing.

## PATENTS

- **Video Summarization Using Semantic Information:** Patent filed in December 2015 (with Intel Labs).
- **Visual Search and Retrieval Using Semantic Information:** Patent filed in May 2016 (with Intel Labs).

## RESEARCH PROJECTS

- **FineGAN: Unsupervised Hierarchical Disentanglement for Fine-Grained Object Generation and Discovery:** We propose FineGAN, a novel unsupervised GAN framework, which disentangles the background, object shape, and object appearance to hierarchically generate images of fine-grained object categories. To disentangle the factors without any supervision, our key idea is to use information theory to associate each factor to a latent code, and to condition the relationships between the codes in a specific way to induce the desired hierarchy. FineGAN achieves the disentanglement for fine-grained classes of birds, dogs, and cars. Using FineGAN's learned features, we also cluster real images as a first attempt at solving the novel problem of unsupervised fine-grained object category discovery. (Advisor: [Prof. Yong Jae Lee](#))
- **Hide-and-Seek: A Data Augmentation Technique for Weakly-Supervised Localization and Beyond:** We propose 'Hide-and-Seek' a general purpose data augmentation technique. The key idea is to hide patches in a training image randomly, in order to force the network to seek other relevant content when the most discriminative content is hidden. The main advantage of Hide-and-Seek over existing data augmentation techniques is its ability to improve object localization accuracy in the weakly-supervised setting. However, it is not tied only to the image localization task, and can generalize to other forms of visual input like videos, as well as other recognition tasks like image classification, temporal action localization, semantic segmentation, emotion recognition, age/gender estimation, and person re-identification. (Advisor: [Prof. Yong Jae Lee](#))
- **Generating High Precision Object Proposals for Weakly-supervised Object Detection:** We propose a novel way of using videos to obtain high precision object proposals for weakly-supervised object detection. We train a weakly-supervised Region Proposal Network (W-RPN) on the discovered video object regions to generate object proposals in weakly-labeled images. Our experiments show that our W-RPN proposals lead to significant improvement in performance of state-of-the-art weakly-supervised object detection approaches. (Advisor: [Prof. Yong Jae Lee](#))
- **Transferring Common-Sense Knowledge for Object Detection:** We propose the idea of transferring common-sense knowledge from source categories to target categories for scalable object detection. Our key idea is to (i) use similarity not at image-level, but rather at region-level, as well as (ii) leverage richer common-sense (based on attribute, spatial, etc.) to guide the algorithm towards the correct detections. (Advisors: [Dr. Santosh Divvala](#), [Prof. Ali Farhadi](#), [Prof. Yong Jae Lee](#))
- **Predicting the Content Diffusion Path in Online Social Networks:** We propose diffusion LSTM, a memory-based deep recurrent network that learns to recursively predict the entire diffusion path of an image through a social network. Evaluated model on Pinterest data, but it is applicable to other image-driven social media as well. (Advisor: [Prof. Yong Jae Lee](#))
- **Identifying First-Person Camera Wearers in Third-Person Videos:** We propose a new semi-siamese convolutional neural network architecture to establish person-level correspondences across first- and third-person videos. We formulate this task as learning a joint embedding space for first- and third-person videos that considers both spatial- and motion-domain cues. (Advisors: [Prof. Yong Jae Lee](#), [Prof. David J. Crandall](#), [Prof. Michael S. Ryoo](#))
- **End-to-End Localization and Ranking for Relative Attributes:** We propose an end-to-end deep network to simultaneously rank and localize relative visual attributes. Unlike previous methods, our network jointly learns the attribute's features, localization, and ranker. We show state-of-the-art ranking results on various relative attribute datasets and our approach is much faster than previous methods. (Advisor: [Prof. Yong Jae Lee](#))
- **Track and Transfer: Watching Videos to Simulate Strong Human Supervision for Weakly-Supervised Object Detection:** We transfer tracked object boxes from weakly-labeled videos to weakly-labeled images to automatically generate

pseudo ground-truth boxes, which replace manually annotated bounding boxes. We obtain state-of-the-art weakly-supervised detection results on the PASCAL 2007 and 2010 datasets. (Advisor: [Prof. Yong Jae Lee](#))

- **Data-Driven Prediction of Egocentric Camera Movement (Master's Thesis):** We record, and analyze, and present to the community, KrishnaCam, a large (7.6 million frames, 70 hours) egocentric video stream along with GPS position, acceleration and body orientation data spanning nine months of the life of a computer vision graduate student. We explore and exploit the inherent redundancies in this rich visual data stream to answer simple scene understanding questions such as: How much novel visual information does the student see each day? Given a single egocentric photograph of a scene, can we predict where the student might walk next? (Advisors: [Prof. Alexei Efros](#), [Prof. Kavvon Fatahalian](#))
- **Predicting Tourist Spots in a City:** In this work we find tourist spot in the city by using Google street view images. We assume tourist spots have discriminative visual elements which are not present in rest of the city. We use mid-level patches based discriminative clustering approach to find these visual elements. We find tourist spots for Paris and Pittsburgh.
- **Object Recognition in Different Illumination Conditions:** Built a model which captures variations in colors due to different illumination conditions while doing object recognition. Used distance from color convex hull (which captures all possible colors of object) as new distance measure for color based template matching. Also pipelined our model with existing Robust Line2D method. (Advisor: [Prof. Martial Hebert](#)) [[Link to Poster](#)]
- **Geometry Directed Browser for Personal Photographs:** Developed a geometry directed photo browser that enables users to browse their personal pictures with the underlying geometry of the space to guide the process. Images are localized to the monument space using state-of-the-art fast localization method. (Advisor: [Prof. P. J. Narayanan](#))
- **Digital Logic Design (DLD) Virtual Labs:** Designed an interactive web tool on simulation of digital logical circuit for the [virtual labs](#) project. It includes adder circuit, multiplexers, decoders, flip-flops, registers, counters etc. Virtual Labs is initiative of Ministry of Human Resources and Development, Gov. of India. Shown in R&D showcase, 2012 at IIIT-H. Currently hosted at this [[link](#)] and used in various educational institutions in India. (Advisor: [Prof. P. J. Narayanan](#))

## COURSE PROJECTS

- **Detecting and Tracking Object Carried by Pedestrian:** Implemented a system to detect and track objects carried by pedestrians in a video stream. Used methods like background subtraction, Kalman filter based tracking, human body symmetrical and periodicity analysis.
- **Correlation Between FMRI and Visual Features:** Constructed a correlation matrix between images using their FMRI responses and tried to find visual features which will give the most similar correlation matrix. We tried low level features like GIST, color histogram, texon and higher level semantic features like object banks and deep features.
- **Fast Nearest Neighbors computation in Videos:** For any given frame, accelerated the computation of nearest neighbors (visually similar frames) in large video collection. Exploited temporal coherency & redundancy present in videos and use an algorithm similar to "[PatchMatch](#)" to compute approximate nearest neighbors.
- **Graphics API Design Like OpenGL:** Developed an OpenGL like graphics API having complete Graphics pipeline functions from modelling, viewing to viewport clipping.
- **Search Engine for Wikipedia:** Developed a mini search engine on ~24 GB Wikipedia dump. Primary (~1.7 GB) and Secondary Index (~65MB) was built on page content, title, category, links and info-box.
- **Classification of the Tweets and Community Detection in Twitter:** Built a text-classification system which classifies tweets into different topical categories. Also detected community and established relation in the Twitter.

## RELEVANT COURSEWORK

- **Graduate Coursework:** Computer Vision, Machine Learning, The Visual World as seen by the Neurons and Machines, Big Data Approaches in Computer Vision, Visual Recognition, Visual Computing Systems, Math Fundamentals for Robotics.
- **Undergraduate Coursework:** Computer Vision, Digital Image Processing, Machine Learning, Statistical Methods in Artificial Intelligence, Artificial Intelligence, Computer Graphics, Information Extraction and Retrieval.

## TECHNICAL SKILLS

Programming/Scripting Languages	C, C++, Python, Java, Lua
Vision/Deep Learning/Graphics	Matlab, OpenCV(moderate), Caffe, Torch, PyTorch, OpenGL
Parallel Programming	CudaC, OpenCL(basic), OpenMP, MKL(basic)

Web development	Web2py, HTML5, CSS, PHP(basic), JSP(basic), Java Applet
Cloud Technologies	Hadoop, EC2
Miscellaneous	FFmpeg, Eclipse, Netbeans, Vim, SVN/Git

## ACADEMIC ACHIEVEMENTS AND AWARDS

- **UC Davis Graduate Student Travel Award**, 2017.
- **Microsoft Azure Research Award**, 2017.
- **AWS Research Grant**, Amazon Web Services, Inc., 2016.
- **Dean's List of Academic Excellence** for all the undergraduate academic years.
- **IIIT-H Research Award**, 2012 – in recognition of research contribution by an undergraduate.
- **Best Poster Award** (Using GPU technologies) from **Nvidia** in Student Research Symposium, **HiPC** (International Conference on High Performance Computing), 2012.
- **All India Rank 1902 (99.8 percentile)** in All India Engineering Entrance Examination (AIEEE) 2009 (**962,119 candidates**).

## SERVICES

- Reviewer, Conference on Computer Vision and Pattern Recognition (CVPR), 2018 and 2019.
- Reviewer, Transactions on Pattern Analysis and Machine Intelligence (TPAMI).
- Reviewer, Indian Conference on Computer Vision, Graphics and Image Processing (ICVGIP), 2014 and 2016.
- Program Committee: 4th Workshop on Egocentric (First-Person) Vision CVPR 2016, International Workshop on Assistive Vision ACCV 2016, International Workshop on Human Activity Analysis with Highly Diverse Cameras WACV 2017, International Workshop on Attention/Intention Understanding ACCV 2018.
- Mentored a high school student for the project “Fruit detection in the images using deep network” as part of Aggiementor program at UC Davis. Student won the first prize at the national level (among 12 universities).