

# KAUSTAV KUNDU

✉ [kkundu10@gmail.com](mailto:kkundu10@gmail.com)

🌐 <https://kaustav-kundu.github.io/>

🔗 [Google Scholar](#)

---

## EMPLOYMENT

### Senior Applied Scientist, AWS AI

October 2019 — Present

- Research Topics

*Multi-modal FMs: Generalization preserving fine-tuning, Orchestration, RLHF.*

*Open set image recognition: Active learning with limited supervision, Self-supervised learning, Backward compabitibility*

- Products Delivered

*Tech lead (TL): Rekognition Moderation API (v6, v6.1, v7)*

*Individual contributor (IC): Rekognition Moderation API (v4, v5)*

*Collaborations: Rekognition Custom Moderation API (v6.1, v7), Titan Image Generator API*

### Applied Scientist, Amazon Go

March 2018 — October 2019

- Research Topics

*Image level: Multi-task learning, Person detection, Semantic Segmentation*

*Video-level: End-to-end real-time action detection of varying action durations from multiple RGB streams*

- Products Delivered

*Individual Contributor (IC): Amazon Go RGB only solution (v2)*

*Tech Lead (TL): End-to-end CV solution for beverage and hot food pick actions*

---

## EDUCATION

### PhD (ABD) in Computer Science, University of Toronto, Canada

January 2014 — December 2017

*Advisors: Raquel Urtasun and Sanja Fidler*

*Thesis Title: Efficient Search Strategies in 3D for Visual Scene Understanding.*

### Masters in Computer Science, Toyota Technological Institute at Chicago, USA

September 2012 — December 2013

*Advisor: Raquel Urtasun*

*Thesis Title: Joint Semantic Segmentation and Depth Prediction in 3D Point Cloud.*

### BTech in Computer Science and Engineering, IIIT Hyderabad, India

August 2008 — May 2012

*Advisor: P J Narayanan*

*Thesis Title: Geometry directed browser for personal photographs.*

---

## INTERNSHIPS

### Research Intern, Apple (SPG)

June 2016 — September 2016

*Mentors: Ruslan Salakhutdinov, Nitish Srivastava, Charlie Tang*

*Project: Lane boundary prediction using Deep Structured Models*

### Research Intern, Apple (SPG)

May 2015 — August 2015

*Mentor: Bart Nabbe*

*Project: Future lane trajectory prediction of vehicles*

---

## RESEACH INTEREST

**Computer Vision and Machine Learning.** Building multi-modal models with limited supervision, which can be used across diverse *in-domain* and *out-of-domain* scenarios and can *reason* with its environment.

---

## AWARDS

- Best paper honorable mention award at **CVPR 2017**
- Outstanding reviewer at **CVPR 2018, CVPR 2021**
- IIIT-H all round achievement award for contribution in cultural, sports and academic life (2012)
- IIIT-H dean's academic list (2008-2012)
- Trophies won in basketball (2009-2012), tennis (2001-2004), volleyball (2010)

---

## PUBLICATIONS

---

### Under Review

- *Holistic Framework for actively learning on large scale open-set image recognition*. Kaustav Kundu, Ketul Shah, Abhay Mittal, Ritwick Chaudhry, Davide Modolo.
- *Towards Omnisupervised Instance Segmentation with Foundation Models*. Arnav Das, Ritwick Chaudhry, Kaustav Kundu, Davide Modolo.
- *PatchML: Patch Based Learning for Multi-label Image Classification*. Lin Zhang, Abhay Mittal, Ritwick Chaudhry, Kaustav Kundu, Davide Modolo.
- *Contrastive Learning for 6D Object Pose Estimation*. Aditya Deshpande, Yuting Wang, Kaustav Kundu, Dongqing Zhang, Onkar Dabeer

### Peer Reviewed

- *Hierarchical Self-supervised Representation Learning for Movie Understanding*. Fanyi Xiao, Kaustav Kundu, Joseph Tighe, Davide Modolo. **CVPR 2022**
- *Id-Free Person Similarity Learning*. Bing Shuai, Xinyu Li, Kaustav Kundu, Joseph Tighe. **CVPR 2022**
- *What to Look at and Where: Semantic and Spatial Refined Transformer for Detecting Human-Object Interactions*. ASM Iftikhar, Hao Chen, Kaustav Kundu, Xinyu Li, Joseph Tighe, Davide Modolo. **CVPR 2022**
- *TubeR: Tubelet Transformer for Video Action Detection*. Zhao et al. **CVPR 2022**
- *Positive-congruent training: Towards regression-free model updates*. Sijie Yan, Yuanjun Xiong, Kaustav Kundu, Shuo Yang, Siqi Deng, Meng Wang, Wei Xia, Stefano Soatto. **CVPR 2021 (Oral)**
- *Exploiting weakly supervised visual patterns to learn from partial annotations*. Kaustav Kundu, Erhan Bas, Michael Lam, Hao Chen, Davide Modolo, Joseph Tighe. **NeurIPS 2020**
- *Pose Estimation for Objects with Rotational Symmetry*. Enric Corona, Kaustav Kundu, Sanja Fidler. **IROS 2018**
- *SurfConv: Bridging 3D and 2D Convolution for RGBD Images*. Hang Chu, Wei-Chiu Ma, Kaustav Kundu, Raquel Urtasun, Sanja Fidler. **CVPR 2018**
- *3D Object Proposals using Stereo Imagery for Accurate Object Class Detection*. Xiaozhi Chen\*, Kaustav Kundu\*, Yukun Zhu, Humin Ma, Sanja Fidler, Raquel Urtasun. **TPAMI 2017**
- *Annotating Object Instances with a Polygon-RNN*. Lluís Castrejón, Kaustav Kundu, Raquel Urtasun, Sanja Fidler. **CVPR 2017 (Best Paper Honorable Mention Award)**
- *Exploiting Semantic Information and Deep Matching for Optical Flow*. Min Bai\*, Wenjie Luo\*, Kaustav Kundu, Raquel Urtasun. **ECCV 2016**
- *Monocular 3D Object Detection for Autonomous Driving*. Xiaozhi Chen, Kaustav Kundu, Ziyu Zhang, Humin Ma, Sanja Fidler, Raquel Urtasun. **CVPR 2016**
- *3D Object Proposals for Accurate Object Class Detection*. Xiaozhi Chen\*, Kaustav Kundu\*, Yukun Zhu, Andrew Berneshawi, Humin Ma, Sanja Fidler, Raquel Urtasun. **NeurIPS 2015**
- *Rent3D: Floor-Plan Priors for Monocular Layout Estimation*. Chenxi Liu\*, Alexander Schwing\*, Kaustav Kundu, Raquel Urtasun, Sanja Fidler. **CVPR 2015 (Oral)**
- *Geometry Directed Browser For Personal Photographs*. Aditya Deshpande, Siddharth Choudhary, P J Narayanan, Krishna Kumar Singh, Kaustav Kundu, Aditya Singh, Apurva Kumar. **ICVGIP 2012 (Oral)**

---

## ACTIVITIES

---

### Mentorship

*Industry Full time*: Yanbei Chen, Rahul Duggal, Chang Liu, Aaditya Singh, Jiarui Cai, Ritwick Chaudhry, Dario Rancati, Abhay Mittal, Yongxin Wang

*Industry Interns*: Arnav Das, Lin Zhang, Ketul Shah, ASM Iftikhar, Sijie Yan, Hengduo Li, Tao Hu, Yujia Chen

*Academia Interns*: Enric Corona

### Conference/Journal Reviewer

*Conferences*: CVPR 2018 - 2024, ECCV 2018 - 2022, ICCV 2019 - 2023, ICLR 2021 - 2023, ICML 2022 - 2024

*Journals*: T-PAMI 2018 - present

### Tutorial/Workshop Organizer

Cross-Model Compatibility in Computer Vision at **ICCV 2021**

### Teaching Assistant

Inference Algorithms and Machine Learning, Intro to ML, Probabilistic Graphical Models, Neural Networks, Intro to Image Understanding, Mathematical Expression and Reasoning for Computer Science, Intro to Visual Computing