ISHAN RAJENDRAKUMAR DAVE (PH.D. CANDIDATE)

Orlando, FL ishandave95@gmail.com GitHub Webpage LinkedIn Google Scholar

Research Interests: Self-supervised Video Representation Learning, Generative AI for Videos, Large Video-Language Foundational models, Video Understanding for Robotics, Privacy-Preserving Computer Vision

Education

Ph.D. in Computer Science University of Central Florida, USA **Aug 2019 – Dec 2024 (Expected)**
Advisor: Dr. Mubarak Shah

B.Tech in Electronics and Communication S.V. National Institute of Technology, India **2013 – 2017**

Work Experience

Apple Inc., Cupertino, USA *PhD AI/ML Intern* **May 2024 – Current**
↪ Video Engineering Group
- Exploring controllability in generative AI

Adobe Inc., San Jose, USA *Research Scientist Intern* **May 2023 – Nov 2023**
↪ Dr. Simon Jenni and Dr. Fabian Caba
- Worked on fine-grained video retrieval from large-scale (millions) of video gallery
- Filed Patent, Paper accepted in ECCV [1]

Adobe Inc., San Jose, USA *Research Scientist Intern* **May 2022 – Nov 2022**
↪ Dr. Simon Jenni
- Developed a novel self-supervised video representation framework by reformulating temporal self-supervision as frame-level recognition tasks and introducing an effective augmentation strategy to mitigate shortcuts.
- Achieved state-of-the-art performance on 10 video understanding benchmarks of linear classification (Kinetics400, HVU, SSv2, Charades), video retrieval (UCF101, HMDB51), and temporal correspondence (CASIA-B).
- Published paper at AAAI [4].

Publications

1. **Ishan Dave**, Fabian Caba, Mubarak Shah, and Simon Jenni. Sync from the Sea: Retrieving Alignable Videos from Large-Scale Datasets, Accepted in European Conference on Computer Vision (**ECCV**), 2024.
2. **Ishan Dave**, Mamshad Nayeem Rizve, and Mubarak Shah. FinePseudo: Improving Pseudo-Labelling through Temporal-Alignablity for Semi-Supervised Fine-Grained Action Recognition. Accepted in European Conference on Computer Vision (**ECCV**), 2024.
3. **Ishan Dave**, Tristan de Blegiers, Chen Chen and Mubarak Shah. CodaMal: Contrastive Domain Adaptation for Malaria Detection in Low-Cost Microscopes, Accepted in 31st IEEE International Conference on Image Processing (**ICIP**), 2024.
4. **Ishan Dave**, Simon Jenni, and Mubarak Shah. No More Shortcuts: Realizing the Potential of Temporal Self-Supervision, AAAI Conference on Artificial Intelligence (**AAAI**), Main Technical Track, 2024.
5. **Ishan Dave**, Mamshad Nayeem Rizve, Chen Chen, and Mubarak Shah. TimeBalance: Temporally-Invariant and Temporally-Distinctive Video Representations for Semi-Supervised Action Recognition, Conference IEEE Computer Vision and Pattern Recognition (**CVPR**), 2023.
6. **Ishan Dave**, Chen Chen, and Mubarak Shah. SPAct: Self-supervised Privacy Preservation for Action Recognition, Conference IEEE Computer Vision and Pattern Recognition (**CVPR**), 2022.
7. **Ishan Dave**, Rohit Gupta, Mamshad Nayeem Rizve, and Mubarak Shah. TCLR: Temporal Contrastive Learning for Video Representation, Computer Vision and Image Understanding (**CVIU**), 2022. **[150+ citations!]**
8. Tristan de Blegiers*, **Ishan Dave***, Adeel Yousaf, and Mubarak Shah. EventTransAct: A video transformer-based framework for Event-camera based action recognition, IEEE/RSJ International Conference on Intelligent Robots and Systems (**IROS**), 2023. (*= equal contribution)
9. Joseph Fioresi, **Ishan Dave**, and Mubarak Shah. TeD-SPAD: Temporal Distinctiveness for Self-supervised Privacy-preservation for video Anomaly Detection, IEEE/CVF International Conference on Computer Vision (**ICCV**), 2023.
10. Tushar Sangam, **Ishan Dave**, Waqas Sultani, and Mubarak Shah. TransVisDrone: Spatio-Temporal Transformer for Vision-based Drone-to-Drone Detection in Aerial Videos. IEEE International Conference on Robotics and Automation (**ICRA**), 2023.
11. **Ishan Dave**, Zacchaeus Scheffer, Akash Kumar, Sarah Shiraz, Yogesh Singh Rawat, Mubarak Shah. GabriellaV2: Towards better generalization in surveillance videos for Action Detection, 4th International Workshop on Human Activity Detection in multi-camera, Continuous, long-duration Video (HADCV'22), at the IEEE Winter Conf. on Applications of Computer Vision (**WACV**), 2022.

12. Mamshad Nayeem Rizve, Ugur Demir, Praveen Tirupattur, Aayush Jung Rana, Kevin Duarte, **Ishan Dave**, Yogesh Singh Rawat, and Mubarak Shah. Gabriella: An online system for real-time activity detection in untrimmed surveillance videos, 25th International Conference on Pattern Recognition (**ICPR**), 2020 **[Best Scientific Paper Award]**

This is a publicly available version of my CV, for articles under-review contact me

Patent

1. Action Recognition System Preserves Privacy in Video Sharing. Researchers: **Ishan Dave**, Mubarak A Shah, Chen Chen. The University of Central Florida. Invention Track Code: 2023-019. (Status: Filed) TechSheet Link

Major Research Projects

Self-supervised Video Representation Learning **May 2020 – present**

- *TCLR Framework*[7] (May 2020- June 2021): Proposed novel temporal contrastive losses to explicitly increase the temporal distinctiveness at two temporal aggregation steps in video tasks: (1) clip-level (2) feature level.
- *TimeBalance Framework*[5] (Aug 2022 - Jan 2023): Studied two complementary self-supervised video representations: (1) Temporally-Invariant (2) Temporally-Distinctive. Proposed a dual teacher-based framework for semi-supervised action recognition using a novel temporal-similarity based reweighting strategy.
- *Mitigating Shortcuts in temporal self-supervision* [4] (May 2022- Jan 2023): Internship work at Adobe.
- *Video Foundational models* (May 2023- Present): Working on improving the visual encoding of Large Video-Language model for the label-efficient Fine-grained Action recognition.

Privacy Preserving Video Understanding **June 2021 – present**

- *Privacy Preserving Action Recognition*[6] (June 2021- Jan 2022): Implemented a privacy-preserving action recognition framework that removes privacy attributes without labels, maintaining competitive performance and achieving best generalization across novel action and privacy attributes.
- *Privacy Preserving Video Anomaly Detection*[9] (Aug 2022- March 2023): Developed a privacy-aware video anomaly detection framework utilizing temporally-distinctive video representations, achieving state-of-the-art tradeoff between privacy protection and utility performance on three popular weakly supervised VAD datasets.
- *Action Fairness* (March 2023- Present): Studying the bias of private attributes (gender, skin color, clothing, etc.) in SOTA action recognition models.

Video Understanding for Robotics **May 2022 – present**

- *TransVisDrone Framework*[10] (May 2022- Jan 2023): Proposed an end-to-end drone detection framework to tackle various challenging real-world scenarios by learning spatio-temporal dependencies of drone motion.
- *EventTransAct Framework*[8] (Aug 2022 - March 2023): Proposed a video transformer-based framework for event-camera based action recognition, which leverages event-contrastive loss and augmentations to adapt the network to event data.
- *Egocentric Multimodal Action Recognition* (May 2023- Sept 2023): Proposed a framework for recognizing actions from egocentric RGB and Depth modalities in an industry-like environment.

Funding Projects

Deep Intermodal Video Analytics (DIVA) program by IARPA *UCF Team Lead* **Sept 2019 – Dec 2021**

- Worked on various aspects of real-world action detection dataset: multi-label correlation, class-imbalance, generalization for unknown facility cameras, improving computational efficiency with knowledge distillation, dealing with noisy data with curriculum learning. [12], [11]
- Lead team UCF and secured first position for consecutive 2 years on target metric competing with other teams from CMU, JHU, UMD, Purdue, IBM, and MIT.

Biometric Recognition and Identification at Altitude and Range (BRIAR) by IARPA **Aug 2022 – Present**

- Worked on a person re-identification project using an adversarial training framework for cloth-change scenarios.

Awards and Honors

Outstanding Reviewer Ranked in the top 2% for review quality among 10,000 reviewers (**CVPR**) **2024**
Nomination The Order of Pegasus Award, University of Central Florida **2024**
1st place Multi-modal Action Recognition challenge (**ICIAP**) **2023**
2nd place, ActivityNet ActEV Challenge (**CVPR**) **2022**
2nd place TRECVID ActEV: Activities in Extended Video **2021**
1st place & Jury Prize, VI-Priors Action Recognition Challenge (**ICCV**) **2021**
1st place, PMiss@0.02tfa, ActivityNet ActEV SDL (**CVPR**) **2021**
1st place, VI-Priors Action Recognition Challenge (**ECCV**) **2020**
1st place, PMiss and nAUDC, ActivityNet ActEV SDL (**CVPR**) **2020**
2nd place, TRECVID ActEV: Activities in Extended Video **2020**

ORCGS Doctoral Fellowship, **2019-2020**
Top 0.5%, Joint Engineering Entrance-Mains exam, India **2013**

Skills

Programming Languages	Python,
Deep learning frameworks	PyTorch, Keras
Tools/Frameworks	OpenCV, SciKit, MATLAB

Coursework

- Advance Computer Vision (*CAP 6412*)
- Advance Machine Learning (*CAP 6614*)
- Computer Vision Systems (*CAP 6411*)

Professional Services

- Mentored students of NSF Research Experience for Undergrad (REU) 2020, 2021 & 2022
- Reviewer of CVPR, ICCV, WACV, TPAMI, TIMM, CVIU, Pattern Recognition, TCSVT, IEEE Access, Multimedia Tools and Application, etc.

Character Referees

Available upon request