

## Core Expertise

Generative Image/Video Diffusion Models, Controllable Generation, Vision-Language Models (VLMs), Fine-Grained Video Understanding, Self-Supervised Learning

## Professional Experience

Adobe Inc., San Jose, CA, USA *Senior Applied Scientist* Jan 2025 – Present  
Adobe Firefly Foundry

- **Lead** controllable image and video generation for enterprise workflows, including large-scale pretraining, supervised fine-tuning (SFT), and adaptation of video diffusion and multimodal foundation models on custom data; **shipped multiple** 2D/3D generation controls to leading animation studios, including **Disney**.
- Train and adapt custom Vision-Language Models (VLMs) and Multimodal Large Language Models (MLLMs) for large-scale data curation and evaluation of image and video generative models.
- **Initiated and led Generative Upscaler** from research proposal and prototype through production model training, quality improvements, and deployment; shipped as **Photoshop's** default Generative Upscale capability, used by 80% of a **15M+ paid user base** and supporting Adobe Firefly Custom Models.
- **Initiated and led Generative Video Refiner / CreativeVR**, driving project direction and end-to-end model development for restoring structural and motion artifacts in AI-generated and real videos; presented at CVPR'26 [1].

Apple Inc., Cupertino, CA, USA *PhD AI/ML Intern* May 2024 – Aug 2024  
Video Engineering Group | Mentor: Dr. Hao Pan

- Enhanced Stable Diffusion for image editing using Vision-Language Models (VLMs) and multimodal foundation models.
- Trained high-resolution diffusion models on a large-scale dataset of 10M samples.

Adobe Inc., San Jose, CA, USA *Research Scientist Intern* May 2023 – Nov 2023  
Mentors: Dr. Simon Jenni and Dr. Fabian Caba

- Built and evaluated a fine-grained video retrieval system over galleries containing millions of videos, extending retrieval beyond coarse semantic similarity to alignable action phases and key events.
- Resulted in a granted US patent and **ECCV 2024 Oral (top 3%)** [6].

Adobe Inc., San Jose, USA *Research Scientist Intern* May 2022 – Nov 2022  
Mentor: Dr. Simon Jenni

- Delivered a self-supervised video representation system and broad evaluation suite spanning video classification, video retrieval, and temporal correspondence across 10 benchmarks [9].

## Selected Technical Contributions

### Multimodal & Video-Language Models

- Enabled **fine-grained video-language retrieval** from a query video and natural-language modification by learning temporally discriminative video representations and contrastively aligning the query with candidate videos [3].
- Reframed **temporal video alignment as a large-scale retrieval problem**, reranking search results with an alignability score and synchronizing the best matches using frame-level representations and DTW for video editing, processing, and understanding [6].
- Learned a **shared embedding space across images, capture time, and geolocation**, using cyclical temporal metric learning to infer when and where an image was captured and enable compositional and text-based image retrieval [4].

### Video Understanding

- Made **video foundation model representations privacy preserving** using a lightweight latent anonymization adapter, reducing privacy leakage while retaining near-baseline utility across action recognition, temporal action detection, and anomaly detection [2].
- Reduced **foreground and background shortcut bias in video action recognition** using adversarial training on static clips with entropy maximization and gradient regularization, improving robustness to spurious visual cues [5].
- Improved **semi-supervised fine-grained action recognition** by learning an action-phase-aware score from DTW-based temporal alignment and using it to refine pseudo-labels when unlabeled data contains novel classes [7].
- Earlier work explored **self-supervised and semi-supervised video representation learning**, including temporal contrastive learning, frame-level temporal self-supervision, complementary temporally invariant/distinctive representations, and privacy-preserving learning across action recognition, retrieval, temporal correspondence, and anomaly detection [12, 10, 9, 11, 14].

## Generative Video & Diffusion

- Built a **diffusion-prior video restoration DiT model** to repair structural and motion artifacts in generated videos, using a lightweight adapter and temporally coherent degradation training with controllable input fidelity [1].

## Robotics & Spatiotemporal Perception

- Adapted a **Video Transformer for event-camera action recognition** using temporal self-attention, event-specific augmentation, and contrastive learning, enabling efficient high-temporal-resolution perception [13].
- Combined object-level features with **VideoSwin spatiotemporal motion modeling** for end-to-end drone-to-drone detection, improving visual perception efficiency and enabling edge deployment for collision and encounter detection [15].

## Education

---

Ph.D. in Computer Science University of Central Florida, USA

Aug 2019 – Dec 2024

- Advisor: Dr. Mubarak Shah
- Dissertation: Towards Label Efficiency and Privacy Preservation in Video Understanding

B.Tech in Electronics and Communication S.V. National Institute of Technology, India

2013 – 2017

## Selected Publications

---

1. **Ishan Dave\***, Tejas Panambur\*, Chongjian Ge, Ersin Yumer, Xue Bai. Diffusion-Prior-Guided Approach for Structure and Motion Restoration in Generative and Real Videos, 2026 (\*= equal contribution) (**CVPR 2026 Workshops**).
2. Joseph Fiorelli, **Ishan Dave**, and Mubarak Shah. Privacy Beyond Pixels: Latent Anonymization for Privacy-Preserving Video Foundational models, International Conference on Learning Representations (**ICLR**), 2026.
3. Animesh Gupta, Jay Parmar, **Ishan Dave**, and Mubarak Shah. From Play to Replay: Composed Video Retrieval for Temporally Fine-Grained Videos. Neural Information Processing Systems (**NeurIPS**), 2025.
4. David Shatwell, **Ishan Dave**, Sirnam Swetha, and Mubarak Shah. GT-Loc: Unifying When and Where in Images Through a Joint Embedding Space. Proceedings of the IEEE/CVF International Conference on Computer Vision (**ICCV**), 2025. **Oral (top 0.6% of papers)**
5. Joseph Fiorelli, **Ishan Dave**, and Mubarak Shah. ALBAR: Adversarial Learning approach to mitigate Biases in Action Recognition, International Conference on Learning Representations (**ICLR**), 2025.
6. **Ishan Dave**, Fabian Caba, Mubarak Shah, and Simon Jenni. Sync from the Sea: Retrieving Alignable Videos from Large-Scale Datasets, European Conference on Computer Vision (**ECCV**), 2024. **Oral (top 3% of papers)**
7. **Ishan Dave**, Mamshad Nayeem Rizve, and Mubarak Shah. FinePseudo: Improving Pseudo-Labeling through Temporal-Alignability for Fine-Grained Action Recognition. European Conference on Computer Vision (**ECCV**), 2024.
8. **Ishan Dave**, Tristan de Blegiers, Chen Chen and Mubarak Shah. CodaMal: Contrastive Domain Adaptation for Malaria Detection in Low-Cost Microscopes, 31st IEEE International Conference on Image Processing (**ICIP**), 2024. **Oral**.
9. **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.
10. **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.
11. **Ishan Dave**, Chen Chen, and Mubarak Shah. SPAct: Self-supervised Privacy Preservation for Action Recognition, Conference IEEE Computer Vision and Pattern Recognition (**CVPR**), 2022.
12. **Ishan Dave**, Rohit Gupta, Mamshad Nayeem Rizve, and Mubarak Shah. TCLR: Temporal Contrastive Learning for Video Representation, Computer Vision and Image Understanding (**CVIU**), 2022. **[280+ citations!]**
13. **Ishan Dave\***, Tristan de Blegiers\*, 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)
14. Joseph Fiorelli, **Ishan Dave**, and Mubarak Shah. TeD-SPAD: Temporal Distinctiveness for Self-supervised Privacy-preserving video Anomaly Detection, IEEE/CVF International Conference on Computer Vision (**ICCV**), 2023.
15. 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.

16. **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.
17. 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**]

## Patents & Selected Recognition

---

### Patents

1. Action Recognition System Preserves Privacy in Video Sharing. Inventors: **Ishan Dave**, Mubarak Shah, Chen Chen. The University of Central Florida. US Patent: US12142053B2. (Status: **Granted**)
2. Identifying and aligning video clips from large-scale video datasets System. Inventors: Simon Jenni, **Ishan Dave**, Fabian Caba, US Patent: US20250342699A1. (Status: **Granted**)

### Awards and Honors

|                                                |                                                                                |           |
|------------------------------------------------|--------------------------------------------------------------------------------|-----------|
| <b>Outstanding Reviewer</b>                    | Ranked in the top 5% for review quality among 17,491 reviewers ( <b>CVPR</b> ) | 2026      |
| <b>Selected for Doctoral Consortium (ECCV)</b> |                                                                                | 2024      |
| <b>Outstanding Reviewer</b>                    | Ranked in the top 2% for review quality among 10,000 reviewers ( <b>CVPR</b> ) | 2024      |
| <b>Nomination</b>                              | The Order of Pegasus Award, University of Central Florida                      | 2024      |
| <b>1<sup>st</sup> place</b>                    | Multi-modal Action Recognition challenge ( <b>ICIAP</b> )                      | 2023      |
| <b>2<sup>nd</sup> place</b>                    | ActivityNet ActEV Challenge ( <b>CVPR</b> )                                    | 2022      |
| <b>2<sup>nd</sup> place</b>                    | TRECVID ActEV: Activities in Extended Video                                    | 2021      |
| <b>1<sup>st</sup> place &amp; Jury Prize</b>   | VI-Priors Action Recognition Challenge ( <b>ICCV</b> )                         | 2021      |
| <b>1<sup>st</sup> place</b>                    | PMiss@0.02tfa, ActivityNet ActEV SDL ( <b>CVPR</b> )                           | 2021      |
| <b>1<sup>st</sup> place</b>                    | VI-Priors Action Recognition Challenge ( <b>ECCV</b> )                         | 2020      |
| <b>1<sup>st</sup> place</b>                    | PMiss and nAUDC, ActivityNet ActEV SDL ( <b>CVPR</b> )                         | 2020      |
| <b>2<sup>nd</sup> place</b>                    | TRECVID ActEV: Activities in Extended Video                                    | 2020      |
| <b>ORCGS Doctoral Fellowship</b>               |                                                                                | 2019–2020 |
| <b>Top 0.5%</b>                                | Joint Engineering Entrance-Mains exam, India                                   | 2013      |

### Professional Services

- Reviewer of CVPR, ECCV, ICCV, WACV, ICRA, IROS, TPAMI, TIMM, CVIU, Pattern Recognition, TCSVT, IEEE Access, Multimedia Tools and Application, etc.
- Program Committee Member of AAAI 2025, 2026.
- Technical Committee, BMVC Privacy and Fairness Workshop 2024.
- Mentored students of NSF Research Experience for Undergrad (REU) 2020, 2021 & 2022.