

Praveen Tirupattur, Ph.D.

MACHINE LEARNING RESEARCHER | COMPUTER VISION | MULTIMODAL AI

praveentirupattur@gmail.com | [LinkedIn](#) | [GitHub](#) | [Google Scholar](#) | [Portfolio](#)

PROFESSIONAL SUMMARY

Machine learning researcher with 8+ years of experience building and deploying computer vision systems across industry and sponsored research. Expertise in image and video understanding, representation learning, multimodal and vision-language models, and production ML. Reduced training compute cost by 30%+, led award-winning real-time video analytics work, and published at ICML, CVPR, AAAI, ICPR, and ACM Multimedia (440+ citations; one U.S. patent).

CORE EXPERTISE

Programming	Python, C++, Java
ML / Computer Vision	PyTorch, TensorFlow, Keras, OpenCV, scikit-learn; image matting, video understanding, action detection, representation learning, multimodal learning, anomaly detection
Generative / Agentic AI	Large language and vision-language models; Amazon Bedrock, AWS Strands Agents, LangChain, Streamlit
Deployment / MLOps	AWS, Google Cloud Platform, Docker, ONNX, REST APIs, Modal, Git
Languages	English, Telugu, Hindi; German (B1)

PROFESSIONAL EXPERIENCE

Machine Learning Researcher | PhotoDay Sep 2024 - Jul 2026

- Owned end-to-end development of production computer vision services for background removal (image matting) and spill correction, spanning dataset curation, model training, evaluation, deployment, monitoring, and maintenance.
- Reduced training compute cost by more than 30% by optimizing data and model-training pipelines, increasing experimentation throughput.
- Optimized inference pipelines to reduce latency and serving cost for revenue-generating photo-processing workflows.
- Improved deployed image-enhancement models for color correction, retouching, and black-and-white generation; curated higher-quality internal datasets to improve robustness and reduce prediction noise.
- Built agentic AI workflows with Amazon Bedrock, AWS Strands Agents, and LangChain to augment customer-support operations for photo studios.

Graduate Research Assistant | University of Central Florida - CRCV, Orlando, FL Aug 2017 - Aug 2024

- Led research on real-time video activity detection for the IARPA DIVA program, developing actor-localization and multi-label action-classification models robust to small objects, scale variation, and class imbalance.
- Built large-scale video preprocessing, training, and evaluation pipelines and improved model efficiency for real-time deployment.
- Led UCF challenge efforts to 1st place in the ActivityNet ActEV SDL 2020 challenge and 2nd place in NIST TRECVID ActEV 2019, competing against leading academic and industry teams.
- Developed skeleton-based gait-recognition models that outperformed RGB baselines through feature-level fusion; led CTTSO project deliverables through contract milestones.
- Contributed to IARPA BRIAR scale-invariant person identification and evaluation pipelines; related work produced a U.S. patent and an ICPR Best Paper Award.

RESEARCH INTERNSHIPS

Research Scientist Intern | Amazon, Palo Alto, CA May 2023 - Aug 2023

- Developed vision-language representation-learning approaches for long-form video understanding.
- Investigated pretrained large language models for improving temporal reasoning in video models.

Research Scientist Intern | Pinterest, Remote May 2022 - Aug 2022

- Developed a unified model for image and video representation learning using large-scale self-supervised multimodal training.
- Outperformed the in-house image-only baseline through multimodal training.

ADDITIONAL EXPERIENCE

Research Assistant & M.Sc. Thesis | DFKI

Feb 2015 - Aug 2016

German Research Center for Artificial Intelligence, Kaiserslautern, Germany

- Built a multimodal violence-detection system using modality-specific SVMs and late fusion.
- Developed Java annotation tooling and REST services integrating myCBR case-based reasoning with object detection for Kognit, a mixed-reality cognitive-assistance system.

Java Developer | XYRALITY GmbH, Hamburg, Germany

Oct 2016 - Jul 2017

- Developed backend game logic for multiplayer strategy games in Java using WebObjects and the Project Wonder framework.
- Implemented and maintained server-side features for multiplayer game systems.

Software Engineer | Progress Software, Hyderabad, India

Aug 2010 - Aug 2013

- Implemented Java and web-product features and resolved defects across enterprise software products.
- Built a proof of concept for a new product-integration use case; received an Employee of the Month award.

SELECTED PUBLICATIONS & PATENT

- **MGD³: Mode-Guided Dataset Distillation Using Diffusion Models** - *ICML*, 2025 (**Oral; top 1%**)
Used a pretrained diffusion model to generate distilled datasets without model retraining or fine-tuning, improving diversity and performance across ImageNet benchmarks.
- **Modeling Multi-Label Action Dependencies for Temporal Action Localization** - *CVPR*, 2021 (**Oral; top 2.5%**)
Introduced an attention-based layer that models simultaneous and sequential action dependencies in untrimmed video.
- **DVANet: Disentangling View and Action Features for Multi-View Action Recognition** - *AAAI*, 2024
Separated action representations from viewpoint-specific cues using transformer queries and supervised contrastive learning.
- **Video Action Detection: Analysing Limitations and Challenges** - *CVPR*, 2022
Analyzed dataset limitations and temporal biases and introduced the MAMA multi-actor, multi-action benchmark.
- **Gabriella: An Online System for Real-Time Activity Detection in Untrimmed Surveillance Videos** - *ICPR*, 2020 (**Best Paper Award**)
Developed an online pipeline for tubelet extraction, activity classification, and temporal merging in surveillance video.
- **Weakly-Supervised Spatiotemporal Anomaly Detection** - *arXiv preprint*, 2026
Applied video-level supervision and multiple-instance ranking loss to localize anomalies across space and time.
- **ThoughtViz: Visualizing Human Thoughts Using Generative Adversarial Network** - *ACM Multimedia*, 2018
Generated visual representations of EEG-encoded thoughts with a conditional generative adversarial network.
- **U.S. Patent 11,468,676** - Methods of Real-Time Spatio-Temporal Activity Detection and Categorization from Untrimmed Video Segments (2022).

EDUCATION

Ph.D., Computer Science | University of Central Florida

2024

Center for Research in Computer Vision (CRCV); Advisor: Prof. Mubarak Shah

M.Sc., Intelligent Systems | Technical University of Kaiserslautern

2016

B.Tech., Computer Science | Jawaharlal Nehru Technological University

2010

AWARDS & PROFESSIONAL LEADERSHIP

- 1st place, ActivityNet ActEV SDL (2020 and 2021); 2nd place, ActivityNet ActEV Challenge (2022); two 2nd-place finishes in TRECVID ActEV.
- Best Paper Award, ICPR; winner, NIST ASAPS Challenge Contest-1; UCF ORC Doctoral Fellowship.
- Organizer, TinyAction ActivityNet Challenge / CVPR workshops (2021 and 2022); NSF REU mentor in 2019, 2020, 2021, and 2024.
- Reviewer for CVPR, ICCV, ECCV, ACM Multimedia, IEEE Transactions on Multimedia, Pattern Recognition, Machine Vision and Applications, and related venues.