





Praveen Tirupattur, Ph.D.

 Mail

 Scholar

 Github

 LinkedIn




 Twitter

 Website





Summary: Computer Vision and Machine Learning researcher with 8+ years of experience across academia and industry, specializing in image processing, video understanding, representation learning, and multimodal learning. Proven track record of delivering production-ready deep learning solutions with publications in top-tier venues including CVPR, ICML, ACM, and AAAI.

Research Interests: Representation Learning, Video Action Understanding, Large Language Models (LLMs), Large Video-Language Foundational Models, Dataset Condensation, Anomaly Detection, GenAI Detection.

Education



Aug 2017 – Aug 2024	 Ph.D. in Computer Science (GPA: 3.8) Center for Research in Computer Vision (CRCV) University of Central Florida, Orlando, Florida, USA Advisor: Prof. Mubarak Shah
Aug 2013 – Aug 2016	 M.Sc. in Intelligent Systems (GPA: 3.8) Technical University of Kaiserslautern, Kaiserslautern, Germany Thesis title: <i>Violence Detection in Videos</i> .
Sep 2006 – Aug 2010	 B.Tech. in Computer Science (GPA: 3.4) Jawaharlal Nehru Technological University, Hyderabad, India

Skills

Programming Languages	 Python, Java, C++
Deep Learning Frameworks	 PyTorch, Keras, Tensorflow
Agentic AI Frameworks	 AWS Strands, AWS Bedrock, LangChain, StreamLit
Cloud Platforms	 Amazon Web Services, Google Cloud Platform
Deployment & MLOps	 Docker, Git, Modal, ONNX, Rest APIs
Libraries & Tools	 Open-CV, scikit-learn, MATLAB
Languages	 English, Telugu, Hindi, German (B1)

Experience

Machine Learning

Sep 2024 – Current	 Machine Learning Researcher at PhotoDay <ul style="list-style-type: none">Developing image post-processing services, including background removal (image matting) and spill correction, using deep-learning models. Responsible for dataset curation, model training, evaluation, and deployment:<ul style="list-style-type: none">Optimized training pipelines, reducing compute costs by over 30% and accelerating model training.Optimized inference pipelines to lower latency and significantly reduce operational costs.Deployed, monitored, and maintained production-grade models, directly contributing to product revenue growth.Curated and maintained high-quality in-house datasets, improving training robustness and reducing prediction noise.Maintained and improved existing computer vision models for photo enhancement and editing tasks, including color correction, retouching and B/W generation.Developing agentic AI solutions to augment internal customer support workflows, improving support effectiveness and contributing to increased sales and revenue for photo studios.
Aug 2017 – Aug 2024	 Graduate Research Assistant at University of Central Florida (CRCV)

Experience (continued)

- Deep Intermodal Video Analytics (**DIVA**) program by IARPA:
 - Worked on solving video activity detection in security videos with large field-of-view. Developed models for actor localization and action classification, handling variations in the scale of objects and class imbalance in a multi-label classification setting.
 - Worked on optimizing data pre-processing pipeline to handle large-scale datasets and enable efficient model training.
 - Led the team at UCF and secured first place in ActEV SDL 2020 challenge (ActivityNet Challenge, CVPR-2020) and second position in TRECVID 2019 challenge while competing with other teams from CMU, JHU, UMD, Purdue, IBM, and MIT.
 - Worked on various aspects of real-time action detection system including, building the data pipeline, improving the computational efficiency of models, and deployment of the system.
- Gait Recognition by CTTSO:
 - Developed skeleton-based gait recognition models, outperforming RGB baselines via feature-level fusion.
 - Successfully led team deliverables, meeting all contract milestones with enhanced performance metrics.
- Biometric Recognition and Identification at Altitude and Range (**BRIAR**) by IARPA:
 - Contributed to the development of a person-identification model aimed at learning robust representations invariant to variations in the scale of individuals resulting from their distance from the camera.
 - Worked on pre-processing the data for model training and building the pipeline for evaluation.

May 2023 – Aug 2023 **Research Intern at Amazon Inc.**

- Worked on representation learning for long-form video understanding with vision-language training.
- Explored the idea of leveraging pre-trained Large Language Models (LLMs) to improve temporal understanding of video models.

May 2022 – Aug 2022 **Research Intern at Pinterest Inc.**

- Worked on building a unified model for both image and video representation learning.
- Explored large-scale self-supervised training to learn representations for multiple visual modalities.
- Obtained improved performance over the in-house image-based model using multi-modal training.

Jan 2016 – Aug 2016 **Master Thesis Student at German Research Center for AI (DFKI)**

- Focused on detecting various types of violent activities from videos using multi-modal features.
- Trained SVM classifiers on each modality and employed late-fusion to detect videos with violence.

Feb 2015 – Aug 2016 **Research Assistant at German Research Center for AI (DFKI)**

- Project - Kognit: Kognit is a tool developed to help dementia patients. It uses cognitive modeling and mixed reality to augment the cognition of the patients.
- Developed a desktop application in Java, to annotate images required for training object detection model.
- Focused on developing a REST-based web-service framework in Java, exposing the functionality of my-CBR (Case-Based Reasoning) system and integrating it with the object detection module.

Software Development

Oct 2016 – July 2017 **Java Developer at XYRALITY GmbH, Hamburg, Germany**

- Worked on developing gaming logic for multi-player strategy games.
- Focused on back-end development working with Wonder frameworks and Web Objects.

Aug 2010 – Aug 2013 **Software Engineer at Progress Software, Hyderabad, India**

- Java and web development involving implementation of various features and bug fixing.
- Built POC for a new use-case involving the integration of existing products.
- Received the best employee of the month award for my efforts.

Research Publications (Citations: 303)









Conference Proceedings

- 1 J. A. Chan-Santiago, **P. Tirupattur**, G. K. Nayak, G. Liu, and M. Shah, “Mgd3: Mode-guided dataset distillation using diffusion models,” in *Proceedings of the International Conference of Machine Learning*, 2025, [Oral Presentation].
- 2 N. Siddiqui, **P. Tirupattur**, and M. Shah, “Dvanet: Disentangling view and action features for multi-view action recognition,” in *Accepted to AAAI Conference*, 2024.
- 3 R. Modi et al., “Video action detection: Analysing limitations and challenges,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 4911–4920.
- 4 M. N. Rizve et al., “Gabriella: An online system for real-time activity detection in untrimmed security videos,” in *2020 25th International Conference on Pattern Recognition (ICPR)*, IEEE, 2021, pp. 4237–4244, [Best Paper Award].
- 5 **P. Tirupattur**, K. Duarte, Y. S. Rawat, and M. Shah, “Modeling multi-label action dependencies for temporal action localization,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021, pp. 1460–1470, [Oral Presentation].
- 6 **P. Tirupattur**, Y. S. Rawat, C. Spampinato, and M. Shah, “Thoughtviz: Visualizing human thoughts using generative adversarial network,” in *Proceedings of the 26th ACM international conference on Multimedia*, 2018, pp. 950–958.




Patents

- 1 Y. S. Rawat, M. Shah, A. J. B. Rana, **P. Tirupattur**, and M. N. Rizve, *Methods of real-time spatio-temporal activity detection and categorization from untrimmed video segments*, US Patent 11,468,676, Oct. 2022.

Awards and Achievements

- 2022  2nd place, ActivityNet ActEV Challenge (CVPR)
- 2021  1st place, ActivityNet ActEV SDL (CVPR)
- 2020  1st place, ActivityNet ActEV SDL (CVPR)
 -  2nd place, TRECVID ActEV: Activities in Extended Video
 -  Best Paper Award at ICPR
 -  Won ASAPS Challenge, Contest-1 (NIST)
- 2019  2nd place TRECVID ActEV: Activities in Extended Video
- 2017  Awarded UCF ORC Doctoral Fellowship

Professional Activities

-  Organized TinyAction ActivityNet Challenge (CVPR 2021, 2022)
-  Mentored students of NSF Research Experience for Undergrad (REU) 2019, 2020, 2021 & 2024
-  Reviewer for CVPR, ICCV, ECCV, CVIP, ACM-MM, IEEE Transaction on Multimedia, Pattern Recognition, Machine Vision and Applications, etc.