Ayush Gupta

443-679-8704 | agupt120@jh.edu |
https://ayush-00.github.io/
Google Scholar

Current Research Interest

Computer Vision, Multi-modal learning, LLMs, Foundational models, Scene Understanding, Domain Adaptation, Gait Recognition, Person Re-ID

Education

Johns Hopkins University — Baltimore, USA
Ph.D. in Computer Science (Advisor: Prof. Rama Chellappa) — *2022 – current*
M.S.E in Computer Science (GPA: 4.0/4.0)

Birla Institute of Technology and Science, Pilani — Pilani, India
B.E. in Computer Science (GPA: 9.58/10) — *2018 – 2022*
Thesis: Temporal self-similarity for Unsupervised Gait Recognition.

Experience

Research Intern — May 2024 – present
SRI International. Mentor: Dr. Anirban Roy

Teaching Assistantship — Aug. 2023 – Dec. 2023
Machine Perception, JHU. Mentor: Prof. Rama Chellappa

Research Assistant — May 2021 – June 2022
CRCV Lab, University of Central Florida, Mentor: Dr. Yogesh S Rawat

Research Intern — May 2020 – July 2020
Indian Institute of Remote Sensing, ISRO. Mentor: Dr. Rekha Anandrao

Publications

- *Ayush Gupta*, Anirban Roy, Rama Chellappa, Nathaniel D. Bastian, Alvaro Velasquez, Susmit Jha, **"VIGOR: Video QA with Temporal Grounding using Open-ended Responses using Multi-modal LLMs"** under submission.
- *Ayush Gupta*, Rama Chellappa **"MimicGait: A Model-Agnostic Approach for Occluded Gait Recognition using Correlational Knowledge Distillation"** WACV 2025.
- Yuxiang Guo, Anshul Shah, Jiang Liu, *Ayush Gupta*, Cheng Peng, Rama Chellappa **"GaitContour: Efficient Gait Recognition based on a Contour-Pose Representation"** WACV 2025.
- *Ayush Gupta*, Rama Chellappa **"You Can Run but not Hide: Improving Gait Recognition with Intrinsic Occlusion Type Awareness"** WACV 2024 **(Oral)**
- Vuong Nguyen, Samiha Mirza, Abdollah Zakeri, *Ayush Gupta*, Rahma Aloui, Khadija Khaldi, Pranav Mantini, Shishir Shah, Fatima Merchant **"Tackling Domain Shifts in Person Re-Identification: A Survey and Analysis"** CVPR 2024 Continual Learning Workshop.
- Basudha Pal, *Ayush Gupta*, Vishal Patel **"EchoSAM: Predicting Ejection Fraction using Segmentation Guided Vision Transformers"** under submission.
- *Ayush Gupta*, Alexander Matasa, Shruti Vyas, Yogesh S Rawat **"GaitZero: Temporal Self-similarity for Unsupervised Gait Recognition"** under submission.
- *Ayush Gupta**, Ashrya Agrawal*, Poonam Goyal, Navneet Goyal **"Visually Guided Knowledge selection for Video Captioning"** under submission.

Projects

Assured Neuro Symbolic Learning and Reasoning (ANSR) DARPA Program

- Developed a framework for reasoning on complex QA pairs with temporal grounding assurance
- Focus on open-vocabulary question answering with temporal grounding abilities in unconstrained videos
- Utilized foundational vision encoders like CLIP and open source LLMs like Mistral and LLaMA

Biometrics Recognition and Identification at Altitude and Range (BRIAR) IARPA program

- Implemented a multi-view gait recognition framework on turbulent data captured from upto 1000m
- Improved gait recognition under occlusion scenarios
- Fusing this approach with other modalities like face and body to identify subjects
- Algorithms integrated into deployable IARPA system pipeline

Undergraduate Thesis: Vision Based Gait Recognition CRCV Lab, University of Central Florida

- Developed approaches for unsupervised gait recognition using RGB datasets like FVG and CASIA-B
- Utilized self-similarity matrices for capturing gait patterns using Transformers
- Implemented unsupervised contrastive learning losses to train the model

Natural Language Video Description Generation ADAPT Lab, BITS Pilani

- Designed a framework for generating natural language descriptions of videos of real scenes
- Utilized external object detectors to extract generalized nouns for the caption
- Used external knowledge bases to supplement the captioning model with specialized versions of the nouns.

CLARIN COVID-19 Disinformation Hackathon LT Group, Universität Hamburg

- Developed models for automatic fact-checking
- Used news crawling APIs and existing datasets like EUvsDisinfo and LIAR Plus to verify a claim.

Landcover Classification using Satellite Imaging IIRS, ISRO

- Used Google Earth Engine to classify satellite image pixels into landcover categories
- Implemented the Spectral Angle Mapper, SVMs and K-Means learning algorithms

Transfer Learning in Semantic Segmentation for Autonomous Vehicles Course Project, Computer Vision

- Collected a dataset, JHUStreet, of street images from a car and pedestrian perspective around Baltimore.
- Trained and evaluated the DeepLabV3 model on the segmentation task on JHUStreet dataset.

Adversarial Attacks and Defences on CNNs Course Project, Machine Perception

- Implemented FGSM, Noise and Carlini Wagner attacks on CNNs
- Implemented Adversarial training to defend against these attacks.

Reviewing Experience

CVPR 2024, NeurIPS 2024, WACV 2025

Awards and Honors

- Merit Scholarship for being in **top 2% of students at BITS Pilani** consistently since Aug. 2018
- Recipient of **DAAD-WISE 2021 scholarship** for a summer project in Universität Hamburg, Germany
- State Rank 1 in National Science and Talent Search Examination (NSTSE) 2017

Technical Skills

Languages: Python, C, Matlab, Java.
Frameworks: PyTorch, Tensorflow, Keras.

Relevant courseworks

Computer Vision, Artificial Intelligence, Machine Perception, Machine Intelligence, Neural Networks and Fuzzy Logic, Digital Image Processing, Data Structures and Algorithms, Computer Programming, Probability and Statistics, Multi-Variable Calculus, Linear Algebra and Complex Analysis, Differential Equations, Database Systems, Object Oriented Programming

Volunteering

Project Lead: Participatory Community Development, Nirmaan Organization May 2019 - Dec. 2019

- Led a team of 10+ members for scouting infrastructural deficiencies in villages nearby Pilani
- Worked on building a rainwater harvesting system and a solar light in Baas Village.