

# Ayush Gupta

443-679-8704 | [agupt120@jh.edu](mailto: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, Person Re-ID

## EDUCATION

---

<b>Johns Hopkins University</b> <i>Ph.D. in Computer Science (Advisor: Prof. Rama Chellappa)</i>	Baltimore, USA 2022 – current
<b>Johns Hopkins University</b> <i>M.S.E in Computer Science</i>	Baltimore, USA 2022 – 2024
<b>Birla Institute of Technology and Science, Pilani</b> <i>B.E. in Computer Science (GPA: 9.58/10)</i>	Pilani, India 2018 – 2022

## EXPERIENCE

---

<b>Research Intern</b> <i>SRI International. Mentor: Dr. Anirban Roy</i>	May 2024 – present
<b>Teaching Assistantship</b> <i>Machine Perception, JHU. Mentor: Prof. Rama Chellappa</i>	Aug. 2023 – Dec. 2023
<b>Research Assistant</b> <i>CRCV Lab, University of Central Florida, Mentor: Dr. Yogesh S Rawat</i>	May 2021 – June 2022
<b>Research Intern</b> <i>Indian Institute of Remote Sensing, ISRO. Mentor: Dr. Rekha Anandrao</i>	May 2020 – July 2020

## PUBLICATIONS

---

- Ayush Gupta, Rama Chellappa “**You Can Run but not Hide: Improving Gait Recognition with Intrinsic Occlusion Type Awareness**” WACV 2024 (Oral)
- 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.
- 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

---

- Grounded visual question answering with Multimodal LLMs** SRI International
- Developed a framework for temporal grounding of complex QA pairs in unconstrained videos
  - Focus on open-vocabulary question answering with temporal grounding abilities
  - 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, JHUSStreet, of street images from a car and pedestrian perspective around Baltimore.
  - Trained and evaluated the DeepLabV3 model on the segmentation task on JHUSStreet 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.