Akshat Sanghvi

B.Tech CSE (Honours) at IIIT Hyderabad
CGPA: 9.01 till 7th Sem
LinkedIn Github Personal Website
Email: akshatsanghvi2021@gmail.com
Contact: +91-6352927215

Synopsis

Final (4th) year B.Tech student in Computer Science at the International Institute of Information Technology Hyderabad (IIIT-H), pursuing an honours degree with a focus on research. Currently working in three research labs here at IIIT-H : CVIT (Computer Vision), RRC (Robotics) and MLL (3D vision). Enjoy playing chess a lot, and served as a coordinator of my college chess club and have a FIDE rating of 1730.

Skills

LANGUAGES
Python, C, C++, JavaScript, HTML/CSS, Bash, x86, cuda

ML FRAMEWORKS
PyTorch, TensorFlow, Numpy

OTHERS
Markdown, Git, Vim, React.js, Flask, Node.js, MySQL, MongoDB

Achievements

2023 Merit List (Monsoon)
2023 Merit List (Spring)
2021 Deans List 2 (Monsoon)
2021 KYPY Rank 186
2020 RMO Qualified
2019 NTSE Stage 2 Qualified
2021 JEE Mains Rank 156
2021 JEE Advanced Rank 2019

Coursework

Computer Vision, Mobile Robotics, Linear Algebra, Digital Image Processing, Advanced NLP, Statistical Methods in AI, Data Structures and Algorithms, Operating Systems and Networks, Information Security, Quantum Information and Computation, Data and Applications

Education

B.TECH. IN CSE *(2021-NOW)*
IIIT - Hyderabad
CGPA : 9.01 (as of 7th sem)

HIGH SCHOOL *(2019-21)*
Green Valley High School
Percentage: 96% (Class 12)

Research Experience

7/23-NOW **Honours (CVIT)** — **Center for Visual Information Technology**

- **Personalized Lip-Reading for Deaf Speakers:** Customized pre-trained Visual Speech Recognition (VSR) models to improve lip-reading performance for out-of-distribution Deaf (and accented) English speakers. Curated a dataset featuring speakers with unclear or no speech to enhance lip-reading accessibility for the Deaf community. (Submitted work currently under review.)
- **Visual Question Answering (VQA) Web App:** Designed and built web applications to showcase the answering capabilities of different VQA models in diverse domains, including Medical, Road, and Document VQA.

Advised by Dr. Jawahar C.V and Dr. Vinay P. Namboodari

10/23-NOW **Independent Study (RRC)** — **Robotics Research Center**

- **Critical Object Estimation for Self-Driving Cars** : Designing and implementing algorithms for real-time path planning in autonomous vehicles, optimizing for computational efficiency by prioritizing essential vehicle interactions, unlike traditional methods that are limited to analyzing a fixed number of closest vehicles.

Advised by Dr. K. Madhava Krishna and Dr. Arun K. Singh

7/24-NOW **Independent Study (MLL)** — **Machine Learning Lab**

- **Compact 3D Scene Representation** Developing methods for 3D scene reconstruction and novel view synthesis using Gaussian Splatting, addressing challenge of large model sizes (up to a gigabyte) for extensive scenes. Leveraging local repetitions and symmetries to achieve significant storage compression without compromising quality.

Advised by Dr. Avinash Sharma and Dr. Charu Sharma

Publications

SEPT. 2024 **DeafVSR: Personalizing Lip Reading for Deaf Speakers**
This work presents a personalized approach to automatic lip reading, considered to be one of the most important assistive technologies for the deaf community. Employed layer-specific fine-tuning to identify the most effective parameters in the pre-trained model for speaker-specific learning. The work is submitted to the prestigious ICASSP conference and is currently under review.

Projects

2023 **Image-Space Manipulation of Objects in Video** — **CV, ML**
Created 2D simulations of object movement in response to virtual forces, analyzing video of tiny motions to infer material properties via modal analysis and a spring-based physics model, predicting pixel reactions to user-defined forces.

2024 **Exemplar Guided Paraphrase Generation** — **NLP, ML**
Developed ML models for paraphrase generation that uses example sentences (exemplars) to guide rephrasing while preserving the original meaning, utilising the concept of contrastive loss on the style feature and content features of the text