

# Arnesh Batra

[arneshbatra2212@gmail.com](mailto:arneshbatra2212@gmail.com) | +91 9315840114 | [linkedin/arnesh-batra-9684a4211](https://www.linkedin.com/in/arnesh-batra-9684a4211) | [github/arnesh2212](https://github.com/arnesh2212)

## EDUCATION

### Bachelor's degree, Computer Science and Artificial Intelligence

| Aug 2023 - May 2027

INDRAPRASTHA INSTITUTE OF INFORMATION TECHNOLOGY, DELHI

Grade : 8.8

## WORK EXPERIENCE

### HMI LAB | RESEARCH INTERN

Delhi, IN | May 2024 – Present

- Conducting research on a project titled "Vision-Audio-Text based Multimodal Modeling and Dataset Generation" under the guidance of Professor **Jainendra Shukla**.
- Implementing Multimodal LLM for various tasks, including video captioning and question-answer generation.
- Applying skills in Deep Learning, LLM, and Linux to enhance project outcomes.

### ROBIFY | AI/ML TEAM MEMBER

Delhi, IN | Jan 2024 - Present

- Developed and simulated self-driving car systems, focusing on both development and integration.
- Utilized advanced tools such as ROS, Gazebo, OpenCV, and PyTorch to optimize project performance and functionality.

## PUBLICATIONS

### Audio Spectrogram Transformer Guided Classification and Information Retrieval For Birds

INTERSPEECH'24

- Developed an audio spectrogram-guided classification framework for accurate bird sound recognition..
- Achieved 100% median performance on F1, Precision, and Sensitivity metrics for 51 bird classes from the Xeno-Canto dataset..

## PROJECTS

### SUPERRESOLUTION FOR GRAVITATIONAL LENSING IMAGES

PYTORCH, DEEP LEARNING

- Worked with various models including GANs, SRResnet, and autoencoder-based denoising.
- Achieved the best AUROC Score of 0.9978 with 98.3% accuracy and SSIM 0.94.
- This repository contains trained models to detect strong lensing images with no substructure, subhalo substructure, and vortex substructure using computer vision models.

### CAPTCHA SOLVER USING CRNN WITH ATTENTION

PYTORCH, DEEP LEARNING, COMPUTER VISION

- Captcha solver with above 96% accuracy made using a Convolutional Recurrent Neural Network with Attention.
- Achieved 4th place in the IIT JODHPUR ML Hackathon.

### 3D RESNET FOR GESTURE CLASSIFICATION

PYTORCH, DEEP LEARNING, VIDEO PROCESSING

- The model achieved 84% accuracy on the validation set on the large Jester Gesture dataset.
- This system will be capable of recognizing a wide range of 3D hand gestures with high accuracy and can be integrated into various applications, enhancing human-computer interaction and user experience.

## SKILLS AND ACHIEVEMENTS

**Languages:** Python, Java, C++, Javascript, HTML, CSS, Flutter, SQL

**Technology:** Deep learning, Machine learning, NLP, CV, Data Structures, Cloud Computing, Git, Linux, Docker

**Libraries:** TensorFlow, Pytorch, numpy, pandas, Scikit-learn, Hugging Face, Langchain, TFJS

**Achievements:** 4th in IIT Jodhpur ML Hackathon, Selected in internal Hackathon For SIH, 5th in IIT Jammu FILL THE VOID() Hackathon