Vaibhav Rathore

M.S by Research
Center For Machine Intelligence and Data Science(C-MInDS)
Indian Institute of Technology Bombay
Github: **github.com/Vaibhavrathore1999**

Email: vaibhav.rathor.in@gmail.com
D.O.B: 13th December 1999
Addr: R.No. C-233, IIT Bombay 400076
Mob.: +91 9634986066
LinkedIn: **vaibhav-rathore-ab696b188**

Degree	University	Institute	Year	CPI/%
MS by Research	IIT Bombay	IIT Bombay	2023-25	9.31
B.Tech	MNNIT Allahabad	MNNIT Allahabad	2018-22	9.10
Intermediate/+2	ISC	De Paul School	2015-17	98.00
Matriculation	ICSE	De Paul School	2015	93.80

PUBLICATION

- **Vaibhav Rathore**, Shubhranil B, Saikat Dutta, Sarthak Mehrotra, Zsolt Kira , Biplab Banerjee. **"When Domain Generalization Meets Generalized Category Discovery."**. Accepted in **CVPR 2025**, Main A* Conference.
- **Vaibhav Rathore**, Divyam Gupta , Biplab Banerjee. **HIDISC*: A Hyperbolic Framework for Domain Generalization with Generalized Category Discovery***. Accepted in **NIPS 2025**, Main A* Conference.
- Anisha Saha , **Vaibhav Rathore**, Abhisek Tiwari , Akash Ghosh , Sai Ruthvik Edara , Sriparna Saha. ***M3 Questioning: Multi-modal, Multi-span Medical Question Answering***. Under Review in **ACM Health**, A* Conference.

WORK EXPERIENCE & RESEARCH INTERNSHIPS

- **Research Intern** | ***Sony Research India***, Bengaluru , India *(May'25 - July'25)*
  Technologies: *PyTorch, Tensorflow, Python, Avatar , 3D Computer Vision , Gaussian Splatting*
  - Designed and implemented a fully automatic 3D lip-syncing pipeline to dub talking-head videos by meshing identity from SPECTRE with audio-driven FLAME parameters from VOCA.
  - Retargeted and blended FLAME lip and jaw coefficients onto SPECTRE's mesh, ensuring topology and temporal consistency for natural articulation.
  - Rendered the animated head back into original footage .
- **Research & Development Intern** | ***Clinical AI Assistance*** *(May'24 - Jul'24)*
  Technologies: *LoRA, Large Language Models (LLMs), Hugging Face Transformers, Streamlit*
  - Fine-tuned **LLMs** with **LoRA**, creating system predicting diseases from 10,000 patient-doctor dialogues
  - Developed **multimodal medical QA model** combining imaging data with text analysis for higher precision
  - Deployed the solution via **Streamlit** for an interactive and efficient diagnostic environment.
- **Graduate Engineer Trainee** | ***Reliance Industries Limited*** *(Aug'22 - Jul'23)*
  Technologies: *SAP , Excel , Data Visualization (Matplotlib/Seaborn)*
  - Built ML models for predictive maintenance and managed shutdown logistics using SAP, reducing downtime and ensuring on-time completion

COMPUTER VISION AND IMAGE ANALYSIS

- **Calorie Estimation of Food Items from Images using Deep Learning** GitHub Link *(Sep'23 - Nov'23)*
  (CS 725: Foundational ML | Instructor : ***Prof. Sunita Sarawagi*** *)*
  - Developed a ML system to estimate **calorie content** from food images for **dietary tracking**.
  - Built a multi-stage pipeline with **YOLO** for food detection & **GrabCut** for precise foreground segmentation.
  - **Ensured accurate analysis and calorie prediction** by isolating food items from complex backgrounds.
- **Fine-Grained Classification on CUB Dataset** GitHub Link *(Mar'24 - May'24)*
  (GNR 638: ML for Remote Sensing | Instructor: ***Prof. Biplab Banerjee****)*
  - **Solved fine-grained visual classification** on the CUB-200 dataset (**200 bird species**), accurately distinguishing subcategories with high visual similarity.
  - Designed a **CNN with <10M parameters**, achieving **35% lower model size** with competitive accuracy.
  - **Attained 87.4% top-1 accuracy**, demonstrating an optimal **trade-off between efficiency and performance**.

- **Medical Image Deblurring** GitHub Link *(Jan'24 - May'24)*
  *(CS 736 : Medical Image Computing | Instructor: **Prof. Suyash P. Awate**)*
  - **Mitigated motion blur** in multi-modal medical images (**30% of scans affected**) to improve diagnostics.
  - **Built a scale-recurrent network** with **spatial-asymmetric attention** for focus on critical regions.
  - **Improved clarity by 24% (PSNR)**, providing higher-quality inputs for reliable medical analysis.

GENERATIVE AI & REPRESENTATIONAL LEARNING

- **Learning with Noisy Labels using Vision Transformer (ViT)** GitHub Link *(Oct'24 - Dec'24)*
  *(GNR 650 : Advanced Deep Learning for Image Analysis | Instructor: **Prof. Biplab Banerjee**)*
  - **Classified images** with **40% label noise** on CIFAR-100, boosting robustness to data corruption.
  - Applied the **state-of-the-art Turtle method** to mitigate mislabeled training data.
  - **Achieved 83% accuracy** with a Vision Transformer, maintaining strong performance despite noise.
- **Autoencoding Beyond Pixels (Million AID Dataset** GitHub Link *(Aug'24 - Nov'24)*
  *(GNR 650 : Advanced Deep Learning for Image Analysis | Instructor: **Prof. Biplab Banerjee**)*
  - Implemented **VAE/GAN** to generate high-fidelity images, improving perceptual score by **18%** over SOTA.
  - Applied a **feature-wise similarity metric** from the GAN discriminator, boosting realism in outputs.
  - Achieved **superior fidelity** and controlled attribute edits via **latent space arithmetic** on **10K+ images**.
- **Zero-Shot Learning** *(Aug'24 - Nov'24)*
  *(GNR 650 : Advanced Deep Learning for Image Analysis | Instructor: **Prof. Biplab Banerjee**)*
  - **Designed zero-shot models** using ViT with DINO pretraining, reaching **0.8%** unseen class accuracy.
  - **Improved generalization** via Near-Instance-Level Attribute Bottleneck, achieving **27.1%** on AWA2.
  - Achieved the highest unseen class accuracy (32.76%) by implementing Class Normalization (CNZSL).

SELF PROJECTS

Kinector: A Text-Conditioned 2D Gesture Generator GitHub Link
- **Generated 2D stick-figure animations** from text using **less than 100-sample** short-video dataset.
- Built pipeline with **MediaPipe** keypoint extraction, **GRU** pose prediction, and text-to-motion animation.
- Used **PyTorch** and Bag-of-Words in an interactive Jupyter notebook for end-to-end workflow.

TECHNICAL SKILLS

Programming Languages	Python, C, C++, Bash
Machine Learning	PyTorch, TensorFlow, OpenCV, Numpy , Pandas, Huggingface
Technologies	Docker, Nano, Git, Vim, LaTeX

POSITIONS OF RESPONSIBILITY & EXTRACURRICULAR ACTIVITIES

- **Teaching Assistant , IIT Bombay**
  - e-PGD: Introduction to Python Programming (Prof.Abir De) *(Jan'25 - Present)*
  - DS 303: Intro to Machine Learning (Prof.Prashanth L.A.) *(Jul'24 - Present)*
  - ME 781: Statistical Machine Learning and Data Mining (Prof.Alankar Alankar) *(Jan'24 - Present)*
- **Student Mentor, MNNIT Allahabad** *(Jul'21 - May'22)*
  - Mentoring group of **20+ freshmen**, academic and social support for successful transition into university life.
  - Acted as bridge between students and faculty, organizing academic resources for first-year challenges.
- **Sports Captaincy & Participation**
  - Led team strategies and coordinated practice sessions for university-level cricket and football teams, contributing to multiple tournament wins.

MACHINE LEARNING / DEEP LEARNING COURSES

- **CS 725: Foundations of Machine Learning**
- **GNR 650: Advanced Topics in Deep Learning**
- **CS 736: Medical Image Computing**
- **CS 601: Algorithms and Complexity**
- **GNR638: Machine Learning for Remote Sensing - II**
- **EE 601: Statistical Signal Analysis**
- **EE 610: Image Processing**
- **SC 607: Optimisation**

ACHIEVEMENTS

- Achieved a **top 1 percentile** rank in the **JEE Main 2018**, securing a position among 1.2M candidates nationwide.
- Attained a **top 1 percentile** rank in the **GATE 2022** examination in both **Mechanical Engineering (ME)** and **Engineering Sciences (XE)**, competing against over 100,000 candidates.