

Shivam Tripathi

✉ shivamt@iisc.ac.in [shivamt-tr.github.io](https://github.com/shivamt-tr) [in shivamtripathi28](https://www.linkedin.com/in/shivamtripathi28) [shivamt-tr](https://github.com/shivamt-tr)

Research Interests

I work on evaluating and improving the physical grounding of video generation models, evaluating how well these models understand and simulate real-world dynamics. I am also interested in multi-modal generation, editing, and personalization.

Education

(*)-Academic Excellence, (†)-I-Div Hons

Indian Institute of Technology, Kanpur

July 2021 - May 2025

MS (Research) in Computer Science

GPA: 9.5*/10

Advisor: Prof. Gaurav Sharma [🔗](#)

Pranveer Singh Institute of Technology, Kanpur

July 2016 - July 2020

B.Tech in Computer Science

GPA: 7.66†/10

Publications

Principia: Relational Physics Tests for Video Models

under review

Varun Varma Thozhiyoor, **Shivam Tripathi**, R. Venkatesh Babu, Anand Bhattad

Objects in Generated Videos Are Slower Than They Appear

CVPR Findings 2026

Varun Varma Thozhiyoor, **Shivam Tripathi**, R. Venkatesh Babu, Anand Bhattad

[arXiv:2512.02016](https://arxiv.org/abs/2512.02016) [🔗](#)

Research Experience

Vision & AI Lab (VAL), Indian Institute of Science, Bangalore

Bangalore, India

Project Assistant

May 2025 – Present

- Advisor: Prof. R. Venkatesh Babu [🔗](#)
- Developed a calibration-free relational consistency paradigm to evaluate foundation video models (Wan, Cosmos, Veo), exposing "sub-Earth" gravity in ball drops; performed LoRA fine-tuning via synthetic data (Blender) to correct this defect and generalize to complex multi-object scenes.
- Curated a comprehensive benchmarking framework, "Principia", for evaluating relational Newtonian physics across seven scenarios (restitution, projectile, friction, rotational inertia, spring-mass, pendulum, and momentum transfer), consisting of an in-house dataset of 350+ videos that can be used to systematically diagnose physical inconsistencies.

Audio-Guided Facial Expression Editing

Kanpur, India

MS Thesis, IIT Kanpur

Jan 2022 – May 2024

- Advisor: Prof. Gaurav Sharma [🔗](#)
- Developed a StyleGAN-based facial expression editing framework, featuring a hierarchical VQ-VAE encoder to invert images to W+ space, and implemented latent-code editing pipeline for identity-preserving expression transfer.
- Built a data curation pipeline using MTCNN and YAMNet to extract and balance 11,863 face-audio pairs from AudioSet, driving cross-modal alignment via InfoNCE loss to map audio signals directly into a facial expression embedding space.

TensorTour (acquired by Typeface.ai [🔗](#))

Remote

Research Intern

May 2022 - July 2022

- Image Retrieval System: Managed image metadata using SQLite, performed CRUD operations, and built a Flask API for content-based image retrieval, providing top-k similar images from user queries
- Explored and conducted a comparative study of available neural image compression models

Indian Statistical Institute, Kolkata

Research Intern

Kolkata, India

Jan 2019 - Sep 2019

- Advisor: Prof. Nikhil R. Pal [🔗](#) ; Internship Letter: [🔗](#)
- Unsupervised Feature Selection: Conducted experiments leveraging self-organizing maps (SOMs) for 2D lattice projection with Sammon's structure-preserving loss, selecting significant features while preserving lattice visualization
- Manifold Learning for Data Visualization: Experimented on t-distributed stochastic neighbor embedding (t-SNE) and autoencoder-based latent representation methods to enhance visualization for datasets with complex manifolds

Professional Experience

Samsung R&D Institute Bangalore, India

Senior Engineer (Research)

Bangalore, India

July 2024 - April 2025

- Part of the Visual Intelligence Team tasked with the development of a proof-of-concept (PoC) for an A-grade patent on 'Image Panning Generation with Mobile Camera.'
- Developed an automated image panning generator for pre-recorded videos, ensuring that the salient object remains in focus while dynamically applying motion blur to the background.
- Implemented object tracking and frame fusion techniques to accurately select panning targets and seamlessly blend frames for creating the panning effect.

Projects

Image Colorization with conditional GANs

[github](#) [🔗](#)

Mentor: Prof. Priyanka Bagade

- Worked on pix2pix image colorization model with deep residual UNet and generator pre-training
- Experimented with generator designs and regularizers; assessed on ImageNet and MSCOCO using PSNR and FID score

Honors and Awards

- **Samsung Professional Level Software Certification** (SWC Professional) for software competency.
- **Academic Excellence Award** [🔗](#) (2021 & 2022) at **IIT Kanpur**, ranking in top 10% GPA.
- **All India Rank 231** in **GATE** (Graduate Aptitude Test) Computer Science 2021 (out of 101,922 candidates).

Teaching Assistant

CS771: Introduction to Machine Learning

Jan 2023 - May 2023

CS300: Technical Communication

July 2022 - Nov 2022

ESC101: Fundamentals of Computing

Dec 2021 - July 2022

Technical Skills and Relevant Coursework

(*)-Awarded Grade 'A'

Languages & Framework: Python (PyTorch), C, C++

Tools & Libraries: OpenCV, Blender, Scikit-Learn, Git

MS Coursework: Introduction to Machine Learning*, Deep Learning for Computer Vision*, Data Mining*, Introduction to IoT