Shivam Tripathi

shivamtr@cse.iitk.ac.in shivamt-tr.github.io shivamtripathi28 shivamt-tr

Research Interests

Representation learning for image, video, 3D, and multi-modal data; creative editing/manipulation using generative AI; working with probabilistic models such as GANs, VAEs, and denoising diffusion models

Education

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

Indian Institute of Technology, Kanpur — *July 2021 - Present*
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

Professional Experience

Samsung R&D Institute Bangalore, India — *Bangalore, India*
Senior Engineer (Research) — *July 2024 – Present*

- 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'
- Developing 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.

Research Experience

Audio-Guided Image Manipulation — *Kanpur, India*
MS Thesis, IIT Kanpur — *Jan 2022 – Present*

- Advisor: Prof. Gaurav Sharma
- Developing an audio-visual stylization framework that modifies image styles based on audio semantics
- Designed a hierarchical VQVAE model to invert input images into StyleGAN's latent space, then perturbed the latent codes using a unified audio-visual feature space to generate stylized results
- Implemented StyleGAN2 inversion, audio-visual feature alignment, and latent code editing
- Build a pipeline to extract training data for audio-visual feature learning utilizing large-scale audio-visual datasets

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 — *Kolkata, India*
Research Intern — *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

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

Analysis of India's Census Data and COVID-19 Data *github*
Mentor: Prof. Arnab Bhattacharya

- Extracted and analyzed data from government-provided APIs for generating insights into COVID-19 trends, identifying peaks for waves, vaccination status, and forecasted dosing milestones
- Analyzed 2011 Census language data, conducting detailed linguistic demography by gender, age, literacy, and geography. Calculated state-wise language distribution, top regional languages, and gender-based distribution of multi-lingual speakers
- Employed Numpy, Pandas, and JSON for wave analysis, vaccination prediction, and forecasting

Safe Vehicle System using Internet of Things *github*
Mentor: Prof. Priyanka Bagade

- Built drowsy driver detection system, utilizing YOLO, Eyenet, and CNNs for face masks, seatbelts, and gaze detection
- End-to-end testing was done on Proteus simulator using Raspberry Pi

Analyzing Various Factors Affecting Climate Change *github*
Mentor: Prof. Arnab Bhattacharya

- Comprehensive analysis of climate change factors: emissions, temperature rise, glacier melt, sea-level rise, plastic, deforestation, linking to disasters and species endangerment; processed 40+ datasets, revealing climate change insights through meticulous data preprocessing and analysis

IoT-based Smart Irrigation System *github*
Mentor: Prof. Priyanka Bagade

- Simulated Arduino Mega 2560-based irrigation system on Wokwi, gathering temperature and humidity from DHT22 sensors; trained neural network to predict water needs, whose flow is controlled via servo motors

Suspicious Activity Detection *github*

- Developed a video classification system using convolutional and recurrent neural networks (CNNs & RNNs) to detect suspicious and safe activities, automating real-time monitoring for enhanced security
- Employed a pre-trained Inception-v3 model for high-level feature extraction from video frames, followed by LSTM for sequence understanding

Scholastic Achievements and Extra-Curricular

- Cleared Samsung Professional Level Software Certification (SWC Professional) for Software Competency
- Served as Department Placement Coordinator at Student's Placement Office, IIT Kanpur; helping the placement team connect with recruiters for the Fall 2023 placements
- Received **Academic Excellence Award** (2021 & 2022) at **IIT Kanpur**, ranking in top 10% GPA
- Ministry of Human Resource Development (MHRD) Assistantship for GATE Qualified Candidates (2021-2023)
- Secured **All India Rank 231** in **GATE CS 2021** out of 101, 922 candidates that appeared for the examination
- Qualified in CodeChef SnackDown 2019 upto Round 1B
- Qualifier in DST (Department of Science and Technology) & Texas Instruments India Innovation Challenge Design Contest 2018, Anchored by IIM, Bangalore

- Participated in **IIT Bombay's e-Yantra Robotics Competition (eYRC-2017)**, reaching Semi-Finals; utilized Arduino IDE, OpenCV, and V-REP for programming, marker detection, and implementing PID controller for path following of a fruit-collector robot
- Certification Course on Machine Learning Specialization at **CloudXLab**

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: C, C++, Python, SQL, LaTeX

Frameworks/Libraries/Tools: PyTorch, TensorFlow, OpenCV, Scikit-Learn, Numpy, Pandas, Git, LaTeX

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

B.Tech Coursework: Data Structures & Algorithms, Operating Systems, Computer Networks, Database Management Systems, Software Engineering, Agile Software Development, Web Technologies, Image Processing*, Data Compression, Artificial Intelligence, Distributed Systems, Data Warehousing & Data Mining