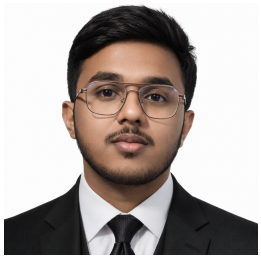
# Soumya Mazumdar

BS (Data Science) – IIT Madras
B.Tech (CSE&BS) – MAKAUT (JIS–GMIT)
Kolkata, West Bengal, India



## Research Focus

**Temporal Generative Modeling:** diffusion for video; long-range temporal consistency
**Sketch/Line-Art Priors:** artist-in-the-loop control; sparse hints; controllable synthesis
**Geometry in Vision:** landmarks/keypoints; head pose; projection-aware conditioning
**Optimization Mindset:** rigorous ablations; reproducible benchmarks; deployment-aware speedups

## Technical Skills

**Programming:** Python, C++, MATLAB
**Frameworks:** PyTorch, TensorFlow, Keras
**CV/Video:** OpenCV, MediaPipe, Decord
**3D/Graphics:** Blender, Open3D, Unity (basic)
**Deep Learning:** Diffusion Models, GANs, VAEs, Vision Transformers, Conformer-CTC
**Systems:** Linux, Git, CUDA, LaTeX
**Math:** Linear Algebra, Probability & Statistics; geometry priors (keypoints/pose)

## Education

**Maulana Abul Kalam Azad University of Technology** *(GMIT campus)*
B.Tech, Computer Science & Business Systems (Hons) Nov 2023 – Jun 2027

**Indian Institute of Technology, Madras**
BS, Data Science and Applications May 2023 – Apr 2027

## Languages

English (Full Professional) Hindi (Full Professional) Bengali (Native)

## Honors & Awards

Swami Vivekananda Merit-cum-Means Scholarship
Kendriya Sainik Board Scholarship

## Links

reachme@soumyamazumdar.com
soumyamazumdar.com
linkedin.com/in/soumyamazumdar
github.com/soumyamazumdar
0009-0006-3521-9557
+91-7278422901

## Research Statement

My research asks: *How can sparse human-authored structure (strokes, line-art, landmarks, pose) be lifted into temporally coherent, controllable video generation?* I engineer geometry-aware generative models that factorize **motion** from **appearance**, then enforce temporal consistency using explicit correspondences (landmarks/pose trajectories, learned point tracks, or flow) and optimization-driven objectives.

Motivated by artist-centric animation pipelines, I am particularly interested in **temporal colorization for sketch/line-art**—conditioning diffusion/transformer models on line structure plus sparse color hints, while preserving edges, identity, and long-range color stability. This aligns with the DynColor direction: assistive colorization that behaves predictably across frames and respects drawn geometry.

## Research Experience

**Variable Energy Cyclotron Centre (DAE, Govt. of India)** — Research Trainee Mar 2025 – Apr 2026

- **Architected** landmark/pose-conditioned video diffusion baselines for talking-head synthesis, coupling 2D facial geometry with temporal attention and identity-preserving conditioning.
- **Engineered** a geometry-conditioning stack (MediaPipe landmarks, head-pose trajectories) and **ablated** guidance strategies for stability vs. controllability across **5 ablation variants (A0–A4)**.
- **Optimized** training/inference via profiling, mixed precision, and reproducible configs; **train time** $\downarrow$ **100%** (train-free pipeline where applicable), **latency** $\downarrow$ **∼90%** (>250 ms $\rightarrow$ 16–33 ms), **throughput** = **30–60 fps** (CPU, $512 \times 512$).
- **Benchmarked** with fidelity + temporal diagnostics; on diffusion baselines, **FID** $\downarrow$ **38.4%** (0.8954 $\rightarrow$ 0.5512) and **FVD** $\downarrow$ **23.8%** (0.6594 $\rightarrow$ 0.5022).
- **Validated** stability improvements in component-wise prototyping: **SyncErr** $\downarrow$ **17.9%** (78 ms $\rightarrow$ 64 ms) with **Low flicker** at **62 fps** (A4).
- **Co-authored** manuscripts/chapters (Springer/Elsevier/CRC/IGI) and **maintained** a clean PyTorch+CUDA training pipeline (versioned configs, deterministic eval, artifact logging).

## Leadership & Roles

**IgniteSpark R&D Cell (GMIT)** — Vice President 2023 – 2025

- **Directed** a 15-member R&D team; **standardized** experiment tracking and peer review across projects.

**Dhyan Chand Sports Council (IIT Madras)** — Eastern Zonal Secretary 2023

## References

- **Vineet Kumar Rakesh** — Senior Scientific Officer, Variable Energy Cyclotron Centre, Department of Atomic Energy, Government of India. Mail: vineet@vecc.gov.in
- **Dr. Hemendra Kumar Pandey** — Senior Scientific Officer, Variable Energy Cyclotron Centre, Department of Atomic Energy, Government of India. Mail: hkpandey@vecc.gov.in
- **Dr. Angshuman Majumdar** — Associate Professor, Brainware University. Mail: hod.ece@brainwareuniversity.ac.in

## Selected Publications

### Journals / Under Review

[1] V. K. Rakesh, **Soumya Mazumdar**, R. P. Maity, S. Pal, A. Das, and T. Samanta, "Advancing Talking Head Generation: A Comprehensive Survey of Multi-Modal Methodologies, Datasets, Evaluation Metrics, and Loss Functions," *The Visual Computer*, vol. 42, no. 9, 2026. doi:10.1007/s00371-025-04232-w.

[2] V. K. Rakesh, **Soumya Mazumdar**, T. Samanta, H. K. Pandey, A. Das, and S. Pal, "Analysis of Hyperparameter Optimization Effects on Lightweight Deep Models for Real-Time Image Classification," *Scientific Reports* (Accepted), 2025. [arXiv:2507.23315].

[3] V. K. Rakesh, **Soumya Mazumdar**, T. Samanta, H. K. Pandey, A. Das, and S. Pal, "Adaptive Video Conferencing Under Severe Bandwidth Constraints Using Audio-Driven Talking-Head Reconstruction," *Engineering Applications of Artificial Intelligence* (under review), 2026.

[4] V. K. Rakesh, **Soumya Mazumdar**, R. P. Maity, S. Pal, A. Das, and T. Samanta, "Quantitative Assessment in Talking Head Generation: A Comprehensive Review of Metrics and Loss Functions," *IET Software* (under review), 2025.

[5] V. K. Rakesh, A. Bhattacharjee, **Soumya Mazumdar**, T. Samanta, H. K. Pandey, A. Das, and S. Pal, "Symbolic Vedic Computation for Low-Resource Talking-Head Generation in Educational Avatars," *IEEE Transactions on Multimedia* (under review), 2026. [arXiv:2602.08775].

[6] V. K. Rakesh, **Soumya Mazumdar**, A. Bhattacharjee, T. Samanta, H. K. Pandey, A. Das, and S. Pal, "Bandwidth-Aware Performance Analysis of Video Conferencing Platforms," *Multimedia Tools and Applications* (under review), 2026.

### Patent

[7] **Soumya Mazumdar**, S. Saha, M. Roy, S. Kundu, P. K. Halder, S. Das, U. Adhikari, and L. Nayak, "Handheld Device for Real-Time Neural Network Analysis," Patent No. 490413–001, Controller General of Patents, Designs & Trade Marks (India), 2026. ipindia.gov.in.

### Conference Proceedings (To Appear)

[8] V. K. Rakesh, **Soumya Mazumdar**, R. P. Maity, T. Samanta, H. K. Pandey, and A. Das, "Empirical Evaluation of State-of-the-Art Talking Head Generation Models," in *Proc. Int. Conf. Recent Advances in Artificial Intelligence & Smart Applications (RAAISA) 2025*, Springer LNNS, to appear (listed as 2026).

[9] V. K. Rakesh, M. Moin, B. Sarkar, **Soumya Mazumdar**, T. Samanta, H. K. Pandey, and A. Das, "THGNLR: An Intelligent Application for Automated, Large-Scale Multimodal Human Audio–Video Data Acquisition and Auto-Cleansing," in *Proc. Int. Conf. AICDAKD 2025*, Springer LNNS, to appear (listed as 2026).

### Book Chapters

[10] **Soumya Mazumdar**, "Educational Implications of 6G Technology for Society 5.0," in *Human-Centric Integration of 6G-Enabled Technologies for Modern Society*, Elsevier, 2025, pp. 255–265. doi:10.1016/b978-0-443-27434-3.00017-9.

[11] **Soumya Mazumdar** and I. Chowdhury, "Integration of Machine Learning, Deep Learning, and Internet of Things Technologies for Stress Management and Identification: A Systematic Analysis," in *Human-Centric Integration of Next-Generation Data Science and Blockchain Technology*, Elsevier, 2025, pp. 407–418. doi:10.1016/b978-0-443-33498-6.00023-6.

[12] H. Hosni and **Soumya Mazumdar**, "6G-Enabled Robotics: The Future of Connectivity in Smart Industrial Manufacturing," in *Industrial Robotics in Smart Manufacturing*, CRC Press (Taylor & Francis), 2025, pp. 241–258. doi:10.1201/9781003614470-12.

[13] **Soumya Mazumdar** and H. Hosni, "Machine Learning and Deep Learning: Necessary Techniques for Extracting Meaningful Information in Large Databases," in *Next Generation Blockchain for Next Generation Society with Futuristic Technologies*, CRC Press (Taylor & Francis), 2026. doi:10.1201/9781003620822-20.

[14] V. K. Rakesh, **Soumya Mazumdar**, R. P. Maity, S. Pal, A. Das, and T. Samanta, "Transparent and Trustworthy Data Practices for Talking Head Generation: A Comprehensive Survey of Datasets and Ethical Implications," in *Centering Transparency and Trust in Data and AI Ecosystems*, IGI Global Scientific Publishing, to appear (listed as 2026).