

HARSH AGARWAL

SENIOR ML ENGINEERING LEAD

Bengaluru, India | harsh.agarwal.ri20@gmail.com | +91-7894405400

PROFILE

Experienced ML leader with 7+ years in building and scaling generative AI systems, **video synthesis technologies, and foundational models**. Proven in leading cross-functional teams and delivering enterprise-grade solutions. Combines deep technical expertise with strategic product thinking. **Founding team member at Rephrase AI (acquired by Adobe)**. [LinkedIn](#) | [Github](#)

EDUCATION

Robotics Institute, Carnegie Mellon University

Master of Science in Robotics, Pittsburgh

CGPA: 4.11/4.33

Aug 2018 to Aug 2020

Indian Institute of Technology - BHU

Bachelor of Technology in Electrical Engineering, Varanasi

CGPA: 9.32/10.00

Jul 2014 to May 2018

PROFESSIONAL EXPERIENCE

Senior Manager, Machine Learning Engineer, Adobe

Firefly

Bengaluru

May 2025 to Present

- **Leading a 16-member Applied ML team** in India, driving roadmap strategy, team expansion, and global collaboration across product, research, and infrastructure to build and scale GenAI systems powering Adobe's creative suite
- Driving research and product development for foundational model for **video editing** using internal Text2Video pipelines and open-source architectures (WAN, Flux).
 - **Designing scalable training and data curation pipelines** supporting both basic (object removal, addition, camera control) and advanced workflows (VFX, stylization, global edits).
 - Collaborating with model teams to **fine-tune Adobe's internal Text2Video models and integrate learnings from open-source research**, improving stability, fidelity, and performance.
 - Developed a PoC for **integrating image referencing into Adobe's internal Text2Video models** (owned design, data, and PoC training).
- Shipped downstream production features including:
 - **Video generation with transparency** — shipped for GA in Firefly.
 - **Object removal in video** — supports 1080p, any duration; integrates SAMv2-based masking and distillation/post-training optimizations, reducing compute cost by ~80%.

Manager Machine Learning Engineer, Adobe

Bengaluru

Nov 2023 to Apr 2025

Led Gen-AI projects within the Digital Video Audio (DVA) group:

- Developed and designed, data curation and training pipeline for **talking-head video generation for dubbing and avatars**.
- **Built a multilingual dataset with 100K speakers across languages, ethnicities, and cultures**.
- Optimized quality and inference for Firefly.com Enterprise, improving the **NPS scores from 2.7 to 3.9 out of 5**, making them **best in market as compared to other competitors**.
- Built video re-targeting pipelines enabling facial expression transfer across identity pairs.

Founding Member and ML Engineer, Rephrase AI

Bengaluru

Jun 2020 to Nov 2023

- Developed **core text-to-avatar systems using CNNs, GANs, and transformers**, reducing training data requirements by 95%, and reducing training time by 70%.
- Built "Rephrase Studio" platform used by Amazon India, Red Ventures, etc.

- Built **celebrity cloning tech powering 100+ personalized marketing campaigns**, generating over 500K videos. Watch [Not Just A Cadbury Ad](#) showcasing personalized avatar generation.
- Owned training infrastructure, actor shoot planning, client delivery, and technical recruitment.

INTERNSHIPS

Research Associate and Teaching Assistant, Prof. Kris Kitani

Aug 2018 to May 2020

Pittsburgh

- Led BLE-based indoor localization achieving high accuracy across multiple phone variants; deployed at AGH Hospital and Pittsburgh International Airport.
- Served as TA for Robotics (16-720B); mentored 200 students, designed course assignments and teaching material.

Research Associate, Australian Centre Robotic Vision

May 2017 to Aug 2017

Adelaide

- Developed unsupervised SLAM systems; achieved SoTA monocular depth/pose estimation.

Research Associate, Video Analytics Lab (VAL), IISc

May 2016 to Aug 2016

Bengaluru

Built segmentation models for hand-drawn sketches using CNNs under Prof. R. Venkatesh Babu.

SKILLS AND FRAMEWORKS

Languages

Python, MATLAB

Frameworks and Libraries

PyTorch, TensorFlow, OpenCV, Hugging Face, Diffusers

Machine Learning

Transformers, Diffusion Models, DiTs, CNNs, GANs, Large Scale Training and fine-tuning

Computer Vision

SLAM, Visual Odometry, Talking Head Faces, Video Retargeting, Video Editing

Management and Leadership Skills

Team Management, Product Strategy, Cross-functional Collaboration, Technical Hiring

PUBLICATIONS

Unsupervised Learning of Monocular Depth Estimation and Visual Odometry with Deep Feature Reconstruction (CVPR 2018)

DeepBLE: Generalizing RSSI-based Localization Across Different Devices