

ROHIT PAUL

AI/ML Engineer — Computer Vision & Generative AI

@ rohitpaul97@gmail.com +91-7577887226 Guwahati, Assam, India
in linkedin.com/in/rohit-paul-91038310b github.com/rohitpaul23
🔗 https://rohitpaul23.github.io/my_portfolio



EDUCATION

Bachelor of Technology

Tezpur University

📅 July 2014 – May 2018 📍 Tezpur, Assam

- Major: Computer Science and Engineering
- GPA: 7.36/10

Master of Technology

IIST, Shibpur

📅 July 2019 – August 2021 📍 Shibpur, West Bengal

- Major: Computer Science and Engineering
- GPA: 9.03/10

Doctor of Philosophy (Discontinued)

Indian Institute of Technology

📅 Jan 2022 – Mar 2022 📍 Delhi

- Major: Computer Science and Engineering

PRACTICAL EXPOSURE

Internship

At NRL — developed a basic web application using the .NET framework.

BTech Project: Semantic Social Networking Site

A social networking blogging website built using AJAX.

DressMe AI

(June 2023 – Oct 2024)

Developed and tested multiple models for a virtual try-on application

- Relit images using a virtual light source with Three.js
- Fine-tuned an image harmonisation model (PHNet)
- Prepared large-scale datasets and trained a virtual try-on model in a distributed system

Alethia AI

(Oct 2024 – Ongoing)

Developed, refined, and optimised a real-time lip-sync model

- Optimised real-time lip-sync pipeline (MoDA) — reduced first-frame latency **from 10s to under 1s**
- Resolved visual artifacts (flickering, wandering eyes, pursed lips) with automated mouth-opening correction for diverse faces
- Built emotive lip-sync with emotion-conditional CFG, intra-emotion interpolation, and Cartesia SSML audio

Created pipelines for on-site generation and asset creation

- Tested and benchmarked SOTA image and video generation models on cost, time, and quality; built deployment APIs for selected models
- Developed prompt templates, a middleware prompt layer, and an agentic AI system for generating assets from requirements — designed to handle noisy user input robustly

RESEARCH & INDEPENDENT PROJECTS

MTech Thesis

(Feb 2020 - May 2021)

VarDenGrid: A New Variable Density Clustering Algorithm with Grid-Based Initialization

An algorithm that works comparably with DB-SCAN and OPTICS, produces variable-density clusters, and requires no user input.

- Accepted at IEEE INDICON 2021 (International Conference of the IEEE India Council)

Other Projects

- TMDB Box Office and House Price Prediction using Random Forests (Kaggle)
- English-to-German NMT model using LSTM networks with attention
- Neural network model for music generation, trained on a music dataset
- WebGL-based real-time lip-sync system and video creation pipeline using Remotion
- Agentic AI workflow using LangChain for automated daily AI news updates with fact check, and auto X and blog post

SKILLS

- **Languages:**
Python | C++ | JavaScript | Java | SQL
- **ML / AI Frameworks:**
PyTorch | TensorFlow | HuggingFace | Diffusers | Sklearn | LangChain | Gradio
- **Inference & Optimisation:**
TensorRT | ONNX | fp16 | DDP | Ray Tune
- **Vision & Media:**
OpenCV | FFmpeg | WebGL | Three.js
- **Cloud & Deployment:**
AWS (S3/SQS/EC2) | GCP | Azure | RunPod | VastAI | Replicate | fal.ai | Docker
- **Dev Tools:**
Git | Linux | FastAPI | Flask | VSCode | Cursor | ClawdBot | ComfyUI

EXTRA ACTIVITIES

- NCC 'A' Certificate
- Venue Coordinator: Tezpur University Technical Event (TECHXETRA)
- Mess Manager Nilachal men's hostel (capacity-416) for the month of October,17
- Solved 115/115 Python challenges in HackerRank scoring 2305 points.