JAY SAWANT

📞 206-850-7888 ✉ jsawant@ucsd.edu in LinkedIn 🎓 Google Scholar 🔗 Jay Sawant

EDUCATION

University of California, San Diego   Sept 2024 – Present
Master of Science in Data Science at HDSI — GPA: **4.0/4.0**   *San Diego, USA*

Indian Institute of Technology Bombay   July 2018 – June 2023
Bachelor and Master of Technology in Electrical Engineering — Major GPA: **8.9/10**   *Mumbai, India*

TECHNICAL SKILLS

Programming: Python (PyTorch, NumPy, Scikit, TensorFlow, Pandas, OpenCV), C++, MATLAB, Bash
Tools & Frameworks: AWS, Slurm, PyTorch-DDP, Linux, Git, Docker, VS Code, Vertex AI, Huggingface
Courses: Deep Generative Models, Reinforcement Learning, Machine Learning, Statistical Models, NLP, Automatic Speech Recognition, Probability and Statistics, Linear Algebra, Medical Image Computing, Data Structures and Algorithms

EXPERIENCE

Synthetic 3D MRI Generation @ Cognitive Imaging Lab, UCSD   Jan 2025 – Present
Graduate Student Researcher   *San Diego, USA*
- Developed a conditional latent diffusion pipeline by integrating a VAE for latent space compression with a 3D diffusion transformer (DiT-3D), enabling the generation of high-fidelity, class-specific synthetic 3D brain MRIs with FID of 16.9
- Augmented the real brain MRI dataset with synthetic data, leading to substantial improvements in F1-score, accuracy, AUC-ROC, and AUCPR for a temporal-lobe epilepsy classifier trained using an EfficientNet-V2 backbone
- Future Objective: Pre-train a 3D ViT on integrated real and synthetic datasets and fine-tune it for downstream tasks

Autonomous Driving Research @ Netradyne   June 2025 – Present
Computer Vision Research Intern   *San Diego, USA*
- Reproduced the TransFuser++ sensor-fusion-based closed-loop validation on the CARLA Leaderboard and Bench2Drive
- Currently prototyping a DINOv2-based image encoder with a lightweight cross-attention for LiDAR-fusion and a novel spatiotemporal attention blocks to capture temporal dynamics to further elevate the closed-loop driving performance

Cardiac Arrhythmia Detection @ Probeplus Solutions   May 2024 – July 2024
AI Consultant   *Bangalore, India*
- Trained a baseline CNN and a multi-head attention-based model as a part of a remote-health monitoring system to achieve a multi-label classification of 26 arrhythmias using a dataset of 12-lead ECG recordings from PhysioNet 2021 challenge
- Improved the model performance by integrating an RNN branch for temporal context and by calculating cross-attention features between CNN and RNN outputs, achieving 2% increase in the challenge score on external test dataset

Deep learning in Histopathology @ MeDAL, IIT Bombay   May 2022 – June 2023
M.Tech Thesis under Prof. Amit Sethi [Publication][M.Tech Thesis]   *Mumbai, India*
- Employed an active learning method to train a classifier, achieving efficient segmentation of WSIs into six tissue regions and outperforming the popular image-processing-based HistoQC tool with higher dice scores on 70% of the WSIs
- Secured 16th place (F1 score 0.67) in the 2023 OCELOT Challenge using DeepLabV3 for cell detection and classification

Contrastive learning in Chest X-rays @ Qure.ai   May 2022 – August 2022
AI Scientist Intern   *Mumbai, India*
- Achieved 0.80 AUC on a 280K-image chest X-ray test set by training ResNet50 on 1.2M scans for opacity classification and boosting performance with a Supervised Contrastive Learning–pretrained backbone

PROJECTS

Person to Person Virtual Try-On | *ECE285: Deep Generative Models, UCSD* [Report]   April 2025 - June 2025
- Generated a paired-person dataset from VITON-HD images using IDM-VITON preprocessing and DensePose
- Engineered a 16-block diffusion-transformer denoiser with self and cross-attention for garment conditioning and FiLM-timestep modulation, optimized via noise-aware EDM, custom 2nd order ODE solver, and CFG to achieve FID 27.7

RAG-based Multimodal Medical QnA App | *MED277: Biomedical NLP, UCSD* [Report]   Oct 2024 - Nov 2024
- Designed a RAG-based multimodal medical chatbot app with multilingual capabilities, integrating a Google Translate API, BiomedBERT embeddings for context retrieval, Gemini 1.5 Pro for output generation, and Gradio for a web interface

Identity Aware Portrait Generation | *CS726: Advanced Machine Learning, IITB* [Report]   Feb 2022 - April 2022
- Adapted CycleGAN for face-to-portrait translation using a FaceNet-based perceptual loss to preserve facial features