Danishjeet Singh

singhdan@iu.edu | singhdan.me | GitHub: danishjeetSingh | LinkedIn: danishjeetSingh

Education

Indiana University, Bloomington — **Bloomington, IN**
Ph.D. in Computer Science — May 2030
M.S. in Computer Science — May 2026
B.S. in Computer Science — May 2025

Experience

Observatory on Social Media — January 2023 - July 2025
Machine Learning Engineer

- Designed and optimized scalable ETL pipelines on Linux servers, leveraging parallelized multiprocessing and Numba JIT compilation to accelerate computation; reduced image similarity analysis across 860 million comparisons from 14 days to 3 minutes.
- Developed deep learning models with PyTorch and TensorFlow, leveraging CLIP and open-source Vision-Language Models (VLMs) such as Qwen 2.5 and Llama 3.2 for deepfake detection and reasoning, achieving over 87% accuracy.

IU Computer Vision Lab, Indiana University — May 2022 - January 2023
Machine Learning Engineer

- Executed large-scale GPU benchmarking experiments on Lambda servers to evaluate optimal hardware configurations for training diffusion models, improving training efficiency for Stable Diffusion, LoRA fine-tuning, ControlNet, and IPA adapters.
- Built a generative AI pipeline leveraging Image Diffusion Models, GANs, and advanced image conditioning techniques to generate 500+ high-fidelity, diverse synthetic human images, ensuring consistency in style and attributes across samples.
- Applied dimensionality reduction (t-SNE, PCA) and clustering (K-Means, DBSCAN) to analyze generative model latent spaces, optimizing hyperparameters for improved representation learning and classifier robustness.

Rite Clinic Private Limited — March 2021 - April 2022
Software Engineer

- Developed a scalable and secure EMR system using React, Node.js, PostgreSQL, and AWS, focusing on frontend usability, backend efficiency, and cloud deployment to ensure seamless performance and compliance with healthcare regulations.
- Implemented key backend functionalities in Node.js and PostgreSQL, designing efficient API endpoints and optimizing database queries to reduce patient record retrieval time by 40%. Integrated JWT authentication and AWS-based storage for handling patient documents securely.

Projects

Deepfake detection on Twitter(https://doi.org/10.54501/jots.v2i4.197)

- Identified 15,000+ deepfake profiles by building a eye-position based heuristic on top of a CNN-based facial detector model, achieving 99% precision and 95% recall across daily active Twitter users.
- Analyzed 10 million profiles using a custom-built image ranking pipeline to systematically assess AI-generated profile likelihood and uncover deepfake usage patterns at scale.

AI-Image detection with Vision-Language Models(https://arxiv.org/pdf/2506.11031)

- Improved detection 29% by developing from scratch a response prefill methodology that steers vision-language models to focus on subtle synthesis artifacts in AI-generated images.
- Validated across multiple models(LLama, Qwen, OpenAI o3) and datasets (faces, objects, animals), demonstrating universal scalability without requiring additional training or fine-tuning.

Denoising Diffusion models(singhdan.me/diffusion)

- Achieved 15% quality improvement by building unconditional and four conditional diffusion models from ground up, incorporating custom EMA and Classifier-Free Guidance implementations for high-quality landscape generation.
- Reached FID score 16.5 by developing attention-based U-Net architecture from scratch on CIFAR-10, enabling controlled class-conditional sample generation with superior visual coherence.

Skills

Languages: Python, R, JavaScript, TypeScript, Java, C#, Bash, JSON, SQL
Databases and Tools: AWS, GCP, MySQL, PostgreSQL, MongoDB, Supabase, Git, Docker, Excel, Snowflake
Data Analysis/Exploration: Pandas, Statistical Models, Matplotlib, Seaborn, Plotly, Regression, Time Series Analysis, DFA, MFDFA
Machine Learning: PyTorch, TensorFlow, NumPy, Scikit-learn, OpenCV, Generative AI, Hugging Face, XGBoost
NLP: Word Embeddings, NLTK, BERT, GPT, Gemini, LLaMA
Development: HTML5, CSS3, React, Next.js, REST API, FastAPI, Flask, Django, Node.js, Express.js, Cursor
Software Engineering: Design Patterns, System Design, Microservices, CI/CD, Distributed Systems