Taneem Ullah Jan

Email: taneemishere@gmail.com
Website: taneemishere.github.io

Applied AI researcher and engineer building scalable systems at the intersection of generative modeling, multimodal learning, and human-centered visual computing.

Education

University of Engineering and Technology Peshawar, Pakistan

Bachelor Studies in Computer Science Sep. 2018 – Sep. 2022

Thesis: HTML Code Generation from Images with Deep Neural Networks
Advisor: Dr. Zakira Inayat
CGPA: 3.58/4.0

Professional Experience

AI Researcher

Sep. 2024 – Present

VOLV AI

- Lead AI research in 3D computer vision and deep generative modeling to deliver a virtual try-on SDK for garments and makeup using neural rendering, GANs, and diffusion models. Adopted by several fashion brands, the SDK increased user engagement by 40%. (VOLV AI Virtual Try-On Demo).
- Drive the development of 2D and 3D digital human avatars with improved pose estimation accuracy using OpenPose and AlphaPose, enabling personalized interactions and reducing product return rates by 30%.
- Spearhead research on multimodal learning systems by integrating large language models (LLMs) with visual and audio signals, developing prototypes that enhance cross-modal reasoning and instruction-following for real-world applications in customer support and interactive experiences.

Research AI Engineer

Jan. 2023 – Dec. 2023

BHuman AI

- Led the development of scalable AI video pipelines using neural image reenactment, facial motion transfer, and voice cloning. Reduced video production costs by 40% for 90K+ users, including major news and media enterprises (Use cases).
- Integrated LLMs (GPT-3.5, LLaMA) with persona avatars and developed fine-tuned RAG-based chatbots for dynamic conversational AI. Achieved 85% user satisfaction across 10+ media and news enterprise clients.
- Researched and implemented state-of-the-art audio-driven neural lip-sync models using GANs and image super-resolution, which became the company's flagship product and primary revenue stream.

Undergraduate Research Assistant

Jan. 2022 – Oct. 2022

CS&IT AI Lab UET Peshawar *Advisor: Dr. Zakira Inayat*

- Developed supervised neural network solutions for image processing and implemented generative models for automated image captioning tasks.
- Contributed to literature reviews and code writing in research projects on contextual intelligence and image similarity. Implemented mathematical optimization techniques like SVD, LU Decomposition, and Cross-Entropy Minimization.

Intern Machine Learning Engineer

Aug. 2021 – Nov. 2021

NAECO Blue GmbH [*web link*]

- Developed analytical functions and machine learning models for selecting optimal spatial and temporal resolution of weather data for solar and wind energy predictions and insights.
- Implemented a data pipeline that reduced research and development time by nearly 50%, which was then adopted by 5+ EU meteorological agencies for their weather prediction models.

Research & Projects

EmbedVoiceLLM: Efficient Multimodal Block-optimized Embedding-Driven Voice Operations with extensible learning

June. 2025 – Sep. 2025

- Engineered a multimodal speech-language system that eliminates ASR pipelines by directly projecting audio embeddings into LLM space, achieving 150ms time-to-first-token with Persistent Adaptive Token strategy and training only 3.5% of parameters through LoRA fine-tuning on Whisper-Llama/NeMo architectures.

FlexiSMPL: Flexible SMPL Body Modeling with Real-time 3D Visualization and Measurement Control

May. 2025 – Aug. 2025

- Built an interactive 3D SMPL body visualization framework with real-time measurement-based control through 23 anthropometric parameters, featuring optimized Open3D rendering with vertex-only updates and 60FPS continuous polling for responsive 3D manipulation and wireframe visualization modes.

VMVLM: Vision-Modulated Vision-Language Models for Improved Instruction Following

Jun. 2025 – Aug. 2025

- Architected a dual-pathway Vision-Language Model combining Q-Former learned queries with direct intermediate ViT feature injection, enabling enhanced multimodal instruction following through complementary visual representations fed to an LLM.

DGM-LLM: Darwin Gödel Machine with Large Language Model Integration for Autonomous Code Self-Improvement

May. 2025 – Jul. 2025

- Engineered an autonomous code optimization system by integrating LLM-guided mutations into an evolutionary algorithm, achieving a 25–35% average improvement across 6 quality metrics through adaptive selection strategies converging in 5–10 generations.

OmniFit-3D: A Unified Framework for 3D Virtual Try-On with Pose-Adaptive Realism

Jan. 2025 – Apr. 2025

- Designed an end-to-end, pose-adaptive 3D try-on pipeline with monocular depth estimation, two-stage clothing warping, and texture fusion, generating realistic meshes from 2D inputs and enabling multi-view rendering on consumer hardware.

LipSyncFace: High-Fidelity Audio-Driven and Lip-Synchronized Talking Face Generation

Jun. 2024 – Jan. 2025

- Developed a two-stage unified network for audio-driven lip-synchronized video synthesis, featuring audio-conditioned sketch prediction at 160×160 resolution and a rendering decoder, achieving PSNR 34.3, LSE-C 7.4, and LSE-D 6.0 with real-time inference capability.

Beyond CNNs: Encoded Context for Image Inpainting with LSTMs and Pixel CNNs

Jan. 2024 – Apr. 2024

- Architected a hybrid image inpainting approach combining WGANs with a novel Row-wise Flat Pixel LSTM architecture that runs efficiently on low-end CPUs, outperforming traditional CNN methods on CIFAR-10 through efficient sequential pixel generation.

lipsync2: Talking Face Generation with Most Accurate Lip Synchronization Aug. 2023 – Dec. 2023

- Enhanced a GAN-based lip-sync framework by incorporating a pre-trained discriminator validation, achieving 10% LSE-C and 6% LSE-D improvements with enhanced long-audio sequence handling.

face2face: One–Shot Talking Head Video Generation from a Source Image Jan. 2023 – Apr. 2023

- Developed one-shot talking head generation using unsupervised motion synthesis with flexible grid-based flow field estimation, achieving 5–10% improvements in animation metrics through adaptive refinement layers.

Publications

Taneem, U. J., & Ayesha, N. (2024). Beyond CNNs: Encoded Context for Image Inpainting with LSTMs and Pixel CNNs. International Journal of Information Systems and Technology (IJIST), 6(5), Special Issue & ICTIS 2024.

Taneem, U. J., & Inayat, Z. (2022). HTML Code Generation from Images with Deep Neural Networks. Journal of Engineering and Applied Sciences (JEAS), UET Peshawar. **(Award: Young Undergraduate Researcher)**

Skills

- **Programming Languages**: Python, C++, MATLAB, LaTeX, Bash Scripting (Linux)
- **ML/AI Frameworks**: PyTorch, PyTorch3D, TensorFlow, Keras, Transformers, LangChain
- **Libraries & Tools**: OpenCV, MediaPipe, ONNX, NumPy, Pandas
- **Developer Tools**: Git, WandB, Docker, Azure AI, GCP (Compute Engine, AI Platform), AWS (EC2, Lambda, SageMaker)