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

- Led AI research in 3D computer vision and deep learning, delivering a virtual try-on SDK for garments and makeup using neural rendering, GANs, and diffusion models. The SDK was adopted by several fashion brands and boosted user engagement by 40% (VOLV AI Virtual Try-On Demo).
- Drove the development of 3D and 2D digital human avatars with improved pose estimation accuracy using OpenPose and AlphaPose, enabling personalized customer interactions and reducing product return rates by 30%.

Research AI Engineer

Jan. 2023 – Dec. 2023

BHuman AI

- Led the development of scalable, personalized AI video pipelines ("1-to-Infinity") using neural image reenactment, animation, facial motion transfer, face-swapping, and voice cloning, which reducing 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, fine-tuned and RAG-based chatbots, and automated AI agents to enable dynamic conversational AI, achieving 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, contributing to 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

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

Jan. 2025 – Present

- Designed a unified framework for 3D virtual try-on that achieves pose-adaptive realism with minimal input requirements and faster execution.
- Integrated 2D image-based virtual try-on with 3D depth estimation to reconstruct 3D try-on meshes, enabling users to visualize clothing fit from multiple perspectives.

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

June 2024 – Jan. 2025

- Developed a two-stage framework for audio-driven talking face generation, addressing the challenge of visual quality and creating a unified training pipeline in lip synchronization.
- Proposed a face generation network to encode visual face information and synchronize lip movements with audio.
- Implemented a rendering decoder to render high-fidelity videos more precisely in lip-sync with faster inference time.

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

- Developed a GAN framework for audio-driven lip synchronization, which generates talking faces.
- Implemented a pre-trained discriminator that evaluates the generated faces for whether audio signals and lips are in sync.
- Achieved 10% improvements on Lip-Sync Error Confidence (LSE-C) and 6% improvements on Lip-Sync Error Distance (LSE-D).

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

- Developed a refinement-based motion transfer method to generate realistic, dynamic talking head videos from a single image, driven by input videos.
- Worked on a pre-trained unsupervised motion synthesis module to estimate hidden motion using flexible grids, addressing the challenge of pose gaps between source and target images.
- Achieved superior performance on benchmarks, exhibiting improvements of $5 - 10\%$ on animation metrics compared to existing approaches.

face–swapping: Swapping Faces in a Video from a Source Image May. 2023 – Aug. 2023

- Developed an image reenactment and image manipulation framework capable of generating realistic face swapped videos from a single image.
- Proposed an identity injection module for transferring the identity information of the source face into the target face at the feature level.

Publications

Taneem, U. J., and Ayesha, N., Beyond CNNs: Encoded Context for Image Inpainting with LSTMs and Pixel CNNs. 2024. ICTIS-24 and IJIST (VOL. 6 NO. 5 Special Issue 2024).

Taneem, U. J., and Zakira, I. HTML Code Generation from Images with Deep Neural Networks. 2022. JEAS UET Peshawar. (**"Young Undergraduate Researcher"** award)

Skills

- **Programming Languages**: Python, C++, MATLAB, MySQL, LaTeX, Bash Scripting (Linux)
- **Frameworks & Packages**: PyTorch, PyTorch3D, TensorFlow, Keras, OpenCV, MediaPipe, LLMs (LangChain), ONNX
- **Developer Tools**: Git, WandB Monitoring Dashboards, Docker, Azure AI, GCP, AWS (*Machine Learning Model Training*)