

# Taneem Ullah Jan

[taneemishere@gmail.com](mailto:taneemishere@gmail.com)  
<https://taneemishere.github.io>

---

## 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

## 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). [\[link\]](#)

**Taneem, U. J.**, and Zakira, I. HTML Code Generation from Images with Deep Neural Networks. 2022. JEAS UET Peshawar. [\[link\]](#)

## PROFESSIONAL EXPERIENCE

### Research AI Engineer

Jan. 2023 – Dec. 2023

[BHuman AI](#)

- Led the development of personalized Generative AI for 1-to-infinity conversational videos. Applied deep learning, computer vision, and graphics to innovate Lip-Sync, Neural Head Avatar, and Image Reenactment algorithms. The goal was to elevate conversational videos by making them natural and personalized.

### Undergraduate Research Assistant

Jan. 2022 – Oct. 2022

*Advisor:* Dr. Zakira Inayat

[CS&IT AI Lab UET Peshawar](#)

- Worked on deep generative models, including transformers for vision and text, and image processing through deep neural networks.
- Conducted studies on mathematical optimization and evaluation techniques to improve the consistency of machine learning models.
- Made notable contributions to research in image similarity and captioning, leading to the successful publication of my thesis in the UET JEAS journal under my advisor's guidance.

### Intern Machine Learning Engineer

Aug. 2021 – Nov. 2021

[\[web link\]](#)

[NAECO Blue GmbH](#)

- Conducted research and evaluation study of intelligent and numerical weather models and their APIs to support the company's data needs. This led to the implementation of a data pipeline that reduced research and development time by nearly 50%.
- Developed analytical functions and machine learning models to enhance the selection of optimal spatial and temporal resolution data for specific locations with greater accuracy.

RESEARCH  
PROJECTS

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

June 2024 – Present

- Developed a two-stage framework for audio-driven talking face generation, addressing the challenge of visual quality in lip synchronization.
- Proposed a face generation network to encode visual face information and synchronize lip movements with audio.
- Enhanced face video quality with a high-resolution decoder, improving realism by 20 – 30% over previous methods with fastest inference time.

**lipsync2: Talking Face Generation with Most Accurate Lip Synchronization**

Aug. 2023 – Dec. 2023

- With the use of neural rendering and GANs, *lipsync2* was focused on generating highly realistic talking head videos with precise lip synchronization.
- By using state-of-the-art mask modeling and lip-sync algorithms, around 26% precision was achieved over *Wav2Lip* in aligning lip movement with associated audio.
- Our ensemble network approach separates masking, alignment, and face generation, ensuring a seamless visual experience.

**face2face: One-Shot Talking Head Video Generation from a Source Image**

Jan. 2023 – Apr. 2023

- Developed a neural 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, to address the challenge of pose gaps between source and target images.
- This method achieved superior performance on benchmarks, exhibiting noticeable improvements of 5 – 10% on animation metrics compared to the existing approaches.

**face-swapping: Swapping Faces in a Video from a Source Image**

May. 2023 – Aug. 2023

- Developed an image reenactment and neural head avatar system capable of generating realistic face swapped videos from a single image.
- This project aimed to develop a framework that transfers the identity of any source face into a target while preserving the target's unique features, such as eye contact and facial expressions.
- Achieved 1080p resolution output videos without using *StyleGAN* or similar external models.

SKILLS

- **Programming Languages:** Python, C++, MATLAB, MySQL, L<sup>A</sup>T<sub>E</sub>X, Bash Scripting
- **Frameworks/Packages:** PyTorch, TensorFlow, Keras, NumPy, OpenCV, Scikit-Learn
- **Developer Tools:** Git, WandB Monitoring Dashboards, GCP, AWS (*Model Training*)
- **Conceptual Topics:** Head Avatars & Digital Humans, Neural Rendering & Synthesis

AWARDS &  
HONORS

- **Young Undergrad Researcher Award** for Bachelor Thesis in Computer Science (2022)
- **Ranked 2<sup>nd</sup>**, BS Computer Science; Batch 18<sup>th</sup> University of Engineering & Technology (2022)
- **Head of the Technical Team** at Google Developer Student Club for two years ('20-21)
- **Ranked 2<sup>nd</sup>**, Intermediate Computer Science, Government College Peshawar Batch 2016<sup>th</sup> ('18)
- Language Ambassador for Pashto at Cohere For AI's AYA Project (2023)