

Muhammad Muaz

✉ m.muaz@utexas.edu • [in Muhammad-Muaz](#) • [Github](#)

Education

M.S. Computer Science

University of Texas at Austin, Austin, Texas

CGPA: 3.88/4.00

Sep 2021-May 2024

Bachelor of Science (BS), Electrical Engineering

Lahore University of Management Sciences (LUMS), Lahore, Pakistan

CGPA: 3.88/4.00

Aug 2017-May 2021

Selected Research Projects

Training Free Efficiency Gains in Video Diffusion Models

PI: Professor Lili Qiu, UT Austin

○ In this project, we analyze and address the computational inefficiencies in video diffusion models via exploiting inherent redundancies in videos.

○ In our approach, we exploited redundancies across inter-frame, intra-frame and motion dimensions in diffusion models similar to video codecs. With our approach, we are able to achieve **25% inference speed up** in open source model (SVD).

○ **Tools used:** Python, Pytorch (Software)

Selected Academic Projects

Conditioning Diffusion Models with 3D Human Shape

Course: *Learning for 3D Humans* – [Link to Report](#)

○ In this project, we explored that previous human image generation (HIG) method are not able to control the generation of human shape and pose with precise control.

○ To answer that, we presented a Control-Net based HIG model that utilizes desired pose and shape as a prior alongside textual input for precise image generation.

○ Both qualitative and quantitative analyses demonstrate that our approach **enhances human shape diversity by 33%** in generated images while maintaining high image quality

○ **Tools used:** Python, Pytorch (Deep Learning)

Comparative Analysis of Multi-Modal Large Language Models (LLM)

Course: *Topics in Natural Language Processing* – [Link to Blogpost](#)

○ Surveyed and compiled three existing ways (query-based ([Instruct BLIP](#)), projection-based ([LLaVA](#)), and parameter-efficient tuning ([LLaMA Adapter](#))) of integrating visual information into LLMs as well as conducted some detailed analysis and comparison among them.

○ Our analysis showed that approaches (i.e., LLaVA) that also finetune the LLM while fusing the additional input (vision for our case study) with the corresponding text showed robustness and better performance to Visual Question-Answering downstream tasks as compared to approaches that freeze LLM and train only the vision module (i.e., InstructBLIP) or train adapters utilizing the visual data (i.e., LLaMA Adapter).

○ **Tools used:** Python, Pytorch (Deep Learning), , Plotly (Visualizations)

Translating Multi-Modal Emotion Recognition Models to Uni-modal, Speech-only counterparts

Course: *Speech Language Technologies* – [Link to Arxiv Submission](#)

○ Innovatively addressed challenges in translating multi-modal emotion recognition model/s to practical and resource-efficient uni-modal speech only counterparts.

○ Proposed a novel framework that leverages knowledge distillation and masked training techniques.

○ Quantitatively validated the effectiveness of both knowledge distillation and masked training in preventing performance drop-off during translation while showcasing that masked training not only prevented drop-off but improved performance over audio-video and audio-only benchmarks.

○ **Tools used:** Python, Pytorch (Deep Learning), Matplotlib (Visualizations)

Social Traversability Estimation in Human Crowded Settings

Course: *Robot Learning* – CS 395T – [Link to Report](#)

- Investigated that classical navigation planners fail to respect social norms in human crowded scenarios.
- Presented a semi-supervised **GAN**-based model(SCNet) for estimating social traversability in navigation paths generated by classical planners using RGB images as input modality.
- Experimental evaluations and analysis indicated that SCNet outperforms CNN-based Neural Network baselines and accurately predicts whether a trajectory is socially compliant in the given scenario.
- **Tools used:** Python, Pytorch (Deep Learning), Matplotlib (Visualizations)

Teaching & Reseach Experience

Graduate Teaching Assistant - UT Austin

- CS-327E(Elements of Databases) Spring 2024
- CS-329E(Elements of Data Visualization) Spring 2023, Fall 2024
- CS-313E(Elements of Software Design) Fall 2022

Graduate Research Assistant - UT Austin

Guide: Professor Lili Qiu

- Conducted initial analysis on the video streaming project via LEO satellite networks by modelling wireless signal propagation between satellite and ground site. May 2021-Aug 2021
- Conducted literature review on using multi-modal data (audio, video, text) for video prediction in context of video streaming and streamlined data collection pipeline for experimentation and model training phase. May 2022-Aug 2023

Scholastic Achievements

- Ranked **5th** overall in undergraduate program among **57** candidates.
- Placed on Dean's Honor List Award for academic years of **2017-18, 2018-19, 2019-20, 2020-21**.
- Received award of **High Distinction** on undergraduate graduation.

Technical Strengths

- **Programming Languages:** C/C++, Python, MATLAB, Bash(Scripting), SQL
- **Deep Learning:** Pytorch, Accelerate, Diffusers, Tensorflow (Keras)
- **Frameworks:** Gradio, Scikit-learn, Pandas, Numpy, OpenCV, L^AT_EX, Git

Coursework

- **Graduate:**
 - **Theory:** Foundational Techniques of Data Science & Machine Learning, Algorithms:Theory & Techniques, Convex Optimization, Advanced Digital Signal Processing, Stochastic Systems
 - **Systems:** Wireless Networking, Digital Communications
 - **Applications:** Machine Learning, Deep Learning, Robot Learning, Speech Processing, Topics in Natural Language Processing, Learning for 3D Humans
 - **Other:** Blockchain & Smart Contract Solutions