

Muhammad Uzair Khattak

muzairkhattak.github.io

Email : muzair.khattak99@gmail.com

LinkedIn: [LinkedIn portfolio](#)

EDUCATION

Master of Science, Computer Vision

Mohamed Bin Zayed University of Artificial Intelligence (MBZUAI), Abu Dhabi

August. 2021 – June 2023

CGPA 4.00/4.00

- Graduate research student supervised by [Dr. Salman Khan](#) and [Dr. Fahad Khan](#)
- MS Research: "Effective transfer of vision-language models with prompt learning for image and video tasks"
- Major courses: Human and Computer Vision, Machine Learning, Visual Object Recognition and Detection

Bachelor of Electrical Engineering

School of Electrical Engineering and Computer Science (SEECs)

National University of Sciences & Technology (NUST), Islamabad, Pakistan

Sept. 2017 – June 2021

CGPA 3.84/4.00

- Undergraduate student supervised by [Dr. Hassan Aqeel Khan](#) and [Dr. Faisal Shafait](#)
- Final year thesis: "Low-Cost Whole Slide Image Scanner with Deep Learning Applications" (Awarded Rector's Gold medal)
- Major courses: Machine Learning, Computer Vision, Signal Processing, Embedded Systems, Microprocessor Systems

RESEARCH INTERESTS

- Multi-modal models for video understanding tasks, zero-shot image recognition and open-vocabulary object detection
- Efficient transfer learning techniques with prompt learning for robustness and generalization of vision-language models
- Reasoning in Video LMMs, Test time optimization, temporal modeling in videos

PUBLICATIONS

* indicates joint first authors, + indicates my role as co-mentor

Learning to Prompt with Text Only Supervision for Vision-Language Models ([Paper](#)) ([Code](#)) Dec 2023 (under review)

Muhammad Uzair Khattak, Muhammad Ferjad Naeem, Muzammal Naseer, Luc Van Gool, Federico Tombari

Align Your Prompts: Test-Time Prompting with Distribution Alignment for Zero-Shot Generalization April 2023 (NIPS-2023)

Jameel Hassan*, Hanan Ghani*, Noor Hussein*, **Muhammad Uzair Khattak**+, Salman Khan, Fahad Khan

Video-FocalNets: Spatio-Temporal Focal Modulation for Video Action Recognition ([Paper](#)) ([Code](#)) March 2023 (ICCV-2023)

Syed Talal Wasim*, **Muhammad Uzair Khattak***, Muzammal Naseer, Salman Khan, Mubarak Shah, Fahad Khan

Learning Self-regulating Prompts for Vision-Language Models ([Paper](#)) ([Code](#)) March 2023 (ICCV-2023)

Muhammad Uzair Khattak*, Syed Talal Wasim*, Muzammal Naseer, Salman Khan, Ming-Hsuan Yang, Fahad Khan

Fine-tuned CLIP models are efficient video learners ([Paper](#)) ([Code](#)) Feb 2023 (CVPR-2023)

Hanoona Rasheed*, **Muhammad Uzair Khattak***, Muhammad Maaz, Salman Khan, Fahad Khan

MaPLe: Multi-modal Prompt Learning ([Paper](#)) ([Code](#)) Feb 2023 (CVPR-2023)

Muhammad Uzair Khattak, Hanoona Rasheed, Muhammad Maaz, Salman Khan, Fahad Khan

Bridging the Gap between Object and Image-level Representations for Open-Vocabulary Detection ([Paper](#)) ([Code](#)) May 2022 (NIPS-2022)

Hanoona Rasheed*, Muhammad Maaz*, **Muhammad Uzair Khattak**, Salman Khan, Fahad Khan

Investigating and Improving Common Loop Closure Failures in Visual SLAM ([Paper](#)) July 2022 (Autonomous Robots)

Saran Khaliq, Muhammad Latif Anjum, Wajahat Hussain, **Muhammad Uzair Khattak**, Momen Rasool

EXPERIENCE

Associate Researcher

Intelligent Visual Analytics Lab (IVAL), MBZUAI, Abu Dhabi

Research topics: Video Large Multimodal Models (LMMs), Prompting techniques for LMMs, Contrastive Vision-Language Models

Mar 2024 – current

- Investigating and bench-marking hallucination in Video LMMs
- Training free-prompting techniques for steering LMM's behavior (self-reasoning)
- Text-only supervision for Vision-Language models with prompt learning
- Co-supervising new master students at IVAL lab

Graduate Research Assistant

Intelligent Visual Analytics Lab (IVAL), MBZUAI, Abu Dhabi

Main research topics: Object detection with transformers, Multi-modal models, Prompt learning

Aug 2021 – July 2023

- Explored CLIP adaptation for action recognition, object recognition and open-vocabulary object detection
- Worked on transformers based detectors including DETR and MDETR for generic and class-agnostic object detection
- Co-mentored new master students at IVAL lab, research outcomes resulted in a NeurIPS'23 paper

- Teaching assistant for Machine Learning course (ML701) and lab instructor for UGRIP internship program at MBZUAI

Undergraduate Research Assistant

May 2019 – May 2021

Signal Processing and Machine Learning (SIGMA) Lab in collaboration with TUKL R&D Center, SEECS, NUST, Islamabad, Pakistan

Main research topics: Medical Imaging, Signal processing, AI on Edge devices, Embedded systems

- Optimization of AI models (YOLOv4, ResNets) with TensorRT on edge devices including Jetson Nano and Jetson TX2
- Real-time auto-focusing with Laplacian filter as focus metric. Integrated algorithm with stepper motors circuitry
- Segmentation of keratin pearl and epithelium tissues using UNET in Whole Slide images
- Participant in international AI competition "Plant Pathology 2020 - FGVC7, CVPR-2020" ([GitHub](#))
- Teaching assistant for Complex Variables and Transforms course (MATH-232)

Artificial Intelligence Intern

Oct. 2020 – April 2021

i engineering Group, Islamabad, Pakistan

- Worked on project "Base station security (BTS) system", created dataset at BTS sites in Pakistan for intruder detection
- Implemented quantized MobileNet model in PyTorch on NVIDIA JETSON Nano 2GB developer kit
- Developed RESTFUL API using Flask and deployed on Heroku web server
- Integrated AI system with i engineering servers and IP camera to provide real time intrusion detection
- Automated the pipeline (updating and shutting down AI algorithm remotely, fetching debug logs) using shell scripts

PROJECTS

Transformers Transforming Vision

Sept 2021 – Dec 2021

- Explored state of the art vision transformers for image recognition including ViTs, DeiTs and T2T-ViTs
- Performance scalability comparison of ViTs and DeiTs with CNNs (ResNets) w.r.t pretraining dataset size
- Benchmarking of ViTs and ResNets on downstream datasets CIFAR10, CIFAR100 and CUB-200 ([GitHub](#))

Understanding the limitations of CNN based pose regressions models for Scene Localization

Oct. 2020 – March 2021

- Explored DL based approaches in Scene localization problem for autonomous robots
- Used custom dataset "1-KM" to understand the drift of CNN pose regressors MapNet and PoseNet in large scenes
- Implementation of visualization software for converting 3-D scene poses to 2-D image projections on RGB images using camera poses and 3D point cloud data

TECHNICAL SKILLS

Languages: Python, C/C++, MATLAB, Assembly and embedded C, SQL, JavaScript, HTML/CSS

Programming: PyTorch, fast.ai, Keras, Latex, TensorRT, Brevitas, Flask, Tkinter, Scikit-learn, Hugging-Face

Tools: Linux, PyCharm, VS code, Git, QuPath, Google Cloud Platform, Heroku, tableau, AUTOCAD

HONOURS AND AWARDS

- MBZUAI graduate studies scholarship holder
- Second [position](#) in RTA Transport Hackathon 2022, organized by Dubai Roads and Transport Authority
- First [position](#) in ADAFSA Agrithon 2021, organized by Abu Dhabi Agriculture and Food Security Authority
- Third position in Hack for Space Hackathon 2021, organized by G42 and coders(hq), UAE
- Winner of the MBZUAI Dogs vs. Cats challenge, 2021 ([report and github repository](#))
- Among the first batch of NVIDIA certified AI Jetson specialists in the world
- Featured in official blog of NVIDIA regarding acknowledgment of AI skills on NVIDIA embedded systems ([Blog](#))
- 3rd time recipient of NUST high achievers award for achieving 4.00/4.00 GPA in three semesters

ACADEMIC SERVICES

- Attended CVPR-2023 conference in Vancouver, Canada.
- Serving as a reviewer at NeurIPS-2023, ECCV-2024, ICCV-2023, ICML-2024, CVPR-2023/2024 & TPAMI
- Served as a volunteer for [Vision Transformers: Theory and applications](#) workshop at ACCV-2022 and NeurIPS-2022
- Attended NeurIPS-2022, CVPR-2022 & NVIDIA GTC 2021 conferences virtually
- Workshop presenter on "Deep Learning using PyTorch", talk organized by IEEE SEECS, NUST ([Recorded session](#))
- Technical trainer at "AI at the Edge" workshop, organized by AI Lounge, Pakistan
- Mentor and project evaluator for new batch of SIGMA lab interns, NUST, Pakistan

EXTRACURRICULAR ACTIVITIES

- Team captain of SEECs, NUST table tennis team
- Mentor at ACM Peer-to-Peer Mentorship program
- Organized "Mall of Humanity" project to distribute clothes to deserving families in the backward areas of Pakistan
- Co-led community service program "Meet the Seniors", aiming to connect high school students with undergraduate students of different fields and universities for face to face mentor-ship sessions

OTHER INTERESTS

- Multiplayer Computer Games, Traveling, Watching movies, Table Tennis, Football, Cricket