

# Muhammad Uzair Khattak

[muzairkhattak.github.io](https://github.com/muzairkhattak)

Email : [muzair.khattak99@gmail.com](mailto: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 – Present

CGPA 4.00/4.00

- Graduate research student supervised by [Dr. Salman Khan](#) and [Dr. Fahad Khan](#)
- MS Thesis: "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"
- Major courses: Machine Learning, Computer Vision, Signal Processing, Embedded Systems, Microprocessor Systems

## PUBLICATIONS

\* indicates joint first authors

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

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

Feb 2023 (CVPR-2023)

### MaPLE: Multi-modal Prompt Learning ([Paper](#)) ([Code](#))

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

Feb 2023 (CVPR-2023)

### Object Detection in Aerial Images: A Case Study on Performance Improvement ([Paper](#)) ([Code](#))

Adnan Khan\*, **Muhammad Uzair Khattak\***, Khaled Dawoud\*

Oct 2022 (ICAIoT 2022)

### Bridging the Gap between Object and Image-level Representations

for Open-Vocabulary Detection ([Paper](#)) ([Code](#))

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

May 2022 (NIPS 2022)

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

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

July 2022 (under-review)

## RESEARCH INTERESTS

- Multi-modal modals 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
- Test time optimization, self-supervised learning

## EXPERIENCE

### Graduate Research Assistant

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

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

Aug 2021 – Present

- 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
- Teaching assistant for Machine Learning course (ML701)

### Undergraduate Research Assistant

Signal Processing and Machine Learning (SIGMA) Lab in collaboration with TUKL

R&D Center, SEECS, NUST, Islamabad, Pakistan

May 2019 – May 2021

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
- Implementation of LeNet Neural Network on ARM-CORTEX M4 discovery board with CMSIS-NN library
- 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

i engineering Group, Islamabad, Pakistan

Oct. 2020 – April 2021

- 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

---

### **Robust generalization of Vision-Language models (on-going)**

Jan 2023 – Present

- Addressing the over-fitting problem of soft prompts during fine-tuning of Vision-Language models
- Designing novel loss formulation and optimization techniques to overcome prompt over-fitting

### **Transformers Transforming Vision**

Sept 2021 – Dec 2021

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

### **Low-Cost Whole Slide Image (WSI) Scanner with Deep Learning Applications**

Nov 2020 – June 2021

- Prototype design of scanning system (RAMPS 1.4 and Arduino Mega) for acquisition of Bio-medical images, using 3D printed structures and electronic circuits. Designed schematics and PCB of hardware circuitry in EAGLE
- Developed Image stitching algorithm using OPENCV, Numpy and scikit-learn. Used cross-correlation and Laplacian blending to acquire [results](#) comparable with open source MS ICE stitching software
- Trained YOLOv4 on BCCD dataset for detection of blood cells in Whole Slide Images. Acquired 94.9% F1-Score on test set
- Comparison of performance metrics by training YOLOv4 on commercial WSI scanner dataset and locally generated dataset (using our scanner) for Mitotic Cells detection

### **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

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

## HONOURS AND AWARDS

---

- MBZUAI graduate studies scholarship holder
- Second [position](#) in Dubai Roads and Transport Authority's (RTA) Transport Hackathon, 2022
- 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](#))
- Rector's gold medal award for best undergraduate thesis, NUST
- 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](#))
- High achievers award recipient from DEAN of SEECS
- Selected for a funded research internship at Technical University of Munich (TUM), Germany for summer 2020

## ACADEMIC SERVICES

---

- Serving as a reviewer at CVPR-2023 & TPAMI Special Issue on *Transformers in Vision*
- Attended NeurIPS-2022, CVPR-2022 & NVIDIA GTC 2021 conference 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

## INTERESTS

---

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