

M. HANAN GANI

🏠 MBZUAI ◊ Masdar City, Abu Dhabi. UAE

☎ (+971)585362287 ☎ (+91)9622517764 ✉ hanan.ghani@mbzuai.ac.ae ◊ m.hanan3829@gmail.com 🌐 [GitHub](#) 🌐 [Homepage](#) 📄 [Google Scholar](#)

EDUCATION

- **Mohamed Bin Zayed University of Artificial Intelligence (MBZUAI)** Abu Dhabi, UAE
Master of Science (MSc.), Machine Learning August 2022 - Present
GPA: 3.78/4.0
Primary Supervisor: [Dr. Salman Khan](#), Associate Professor (✉ Salman.Khan@mbzuai.ac.ae)
Secondary Supervisor: [Dr. Fahad Khan](#), Full Professor (✉ Fahad.Khan@mbzuai.ac.ae)
Mentor and research collaborator: [Dr. Muzammal Naseer](#), Research Scientist (✉ Muzammal.Naseer@mbzuai.ac.ae)
Research Topics: Label-Efficient Learning - Generative models and LLMs; Multi-modal learning
- **National Institute of Technology (NIT)** Srinagar, Kashmir, India
Bachelor of Technology (B.Tech), Electronics and Communication Engineering 2014-2018
Overall GPA: 8.561/10 (Among top 5 of the class)
Supervisors: [Dr. Shahid Mehraj Shah](#) (Assistant Professor, NIT Srinagar, mail: shahidshah@nitsri.net), [Dr. G. R. Begh](#) (Associate Professor, NIT Srinagar, mail: grbegh@nitsri.ac.in)
Undergrad project and thesis: Machine Learning based channel estimation; Real-time Emotion Recognition
- **Saint Joseph's Higher Secondary School** Baramulla, Kashmir (India)
Higher Secondary Part II (Class XII), JKBOSE 2014
Percentage: 96% | Major in Physics, Chemistry, Mathematics and English (Among top 10 of roughly 35k students in the entire J&K state)

WORK EXPERIENCE

- **Mohamed Bin Zayed University of Artificial Intelligence** Abu Dhabi, UAE
Teaching Assistant Sept 2023 - Present
Spring 2024
 - ☐ *Advanced Topics in Vision and Language (CV806)* with [Professor Ivan Laptev](#)
 - ☐ *Deep Learning (AI702)* with [Dr. Haris Khan](#)**Fall 2023**
 - ☐ *Probabilistic and Statistical Inference (ML703)* with [Professor Kun Zhang](#)
 - ☐ *Machine Learning (ML701)* with [Dr. Samuel Horvath](#)Highlights of work: Instruct lab sessions for the courses; grade assignments, exams and quizzes; mentor students for the course projects.
- **Harman International - Connected Car R&D (Samsung)** Bengaluru, India
Machine Learning Research Engineer Oct 2018 - Sept 2021
Subdivision: Machine Learning R&D Team, Harman Connected Car
Highlights of Projects:
 - ☐ Developed **Screen Reliability system** which works real-time and detects anomalies in continuous video streams on HMI screens. Employed deep learning techniques, specifically a Auto-encoders and Generative Adversarial Networks (currently being used in production at Harman facilities).
 - ☐ Developed **Test Case Recommender** which uses transformer based language models to maps user's text query with the relevant test cases to fix automation issues such as software run failures or system crashes. (currently being used in production at Harman facilities)
 - ☐ Developed **Log Failure Categorization** which utilizes error logs to distinguish between software and hardware failures. (currently being used in production at Harman facilities, saves 2 hours per day to software team)
 - ☐ Developed a **Hybrid Icon Detection System** combining the strengths of both classical and deep learning models to detect various buttons on screens for automation testing. (currently being used in production at Harman facilities)
 - ☐ Devised a system, called **Similar Issue Recommender**, which accepts the detailed description of a software issue and uses language model to recommend similar types of software issues fixed in the past. It gives an idea to the user / developer as to what fix could be applied to the issue. (currently being used in production at Harman facilities)

RESEARCH EXPERIENCE





- **King Abdullah University of Science and Technology (KAUST)** Thuwal city, Saudi Arabia
Visiting Student June 2023 - August 2023
Advisor: [Dr. Peter Wonka](#), Full Professor and Associate Director of Visual Computing Center (VCC), CS Department (peter.wonka@kaust.edu.sa)
Highlights of Research:
 - ☐ **Text-to-Image generation from complex and detailed textual prompts**: Diffusion-based generative models encounter challenges when processing lengthy and intricate textual prompts describing complex scenes with multiple objects. We present a novel approach leveraging Large Language Models (LLMs) to extract critical components from textual prompts and use a two-stage mechanism to guide the image generation that aligns with the long textual prompt.

- Mohamed Bin Zayed University of Artificial Intelligence (MBZUAI)** Masdar city, Abu Dhabi, UAE
 Research Assistant - Full time Sep 2021 - Sep 2022
 Senior Advisor: **Dr. Mohammad Yaqub**, Associate Professor at MBZUAI, (mohammad.yaqub@mbzuai.ac.ae)
 Research Collaborations: **Dr. Muzammal Naseer**, Research Scientist, MBZUAI (muzammal.naseer@mbzuai.ac.ae)
 Lab: BiomedIA AI Lab, Computer Vision Department
 Highlights of Research:
 □ **Improving performance of Vision Transformers on small-scale datasets**: We propose a self-supervised weight learning scheme from low-resolution views created on small datasets. This serves as an effective weights initialization to successfully train ViTs from scratch, thus eliminating the need for large-scale pre-training.
- Fatima Fellowship - One year Predoctoral Fellowship in Artificial Intelligence** U.S.A (remote)
 Part-time Fellow April 2021 - Dec 2021
 Mentor: **Dr. Abubakar Abid**, Machine Learning Lead at Hugging Face Inc (USA), Founder at Gradio Inc. (a12d@stanford.edu)
 Highlights of Research: **Multi-Task Learning (MTL)** presents a formidable challenge in deep learning. As part of the Fatima Fellowship, I worked with Dr. Abubakar Abid to advance smart MTL, enabling AI algorithms to handle multiple tasks simultaneously with limited computational resources. Our approach optimizes Vision Transformers (ViTs) by exploiting class-token and self-attention mechanisms, ensuring efficient training of multiple tasks within a constrained computational budget. Selected as **Oral** paper at UAE Graduate Student Research Conference (GSRC). [Project demo code](#)
- Indian Institute of Science (IISc)** Bengaluru, India
 Deep Learning Research Intern Full time: Dec 2017 - Feb 2018 , Part time: March 2018 - June 2018
 Lab: Computational Intelligence & UAV Lab, Aerospace Engineering Department, IISc
 Highlights of Research: Conducted extensive research in Deep Learning and Computer Vision, leading the project **Disguised Facial Recognition using Deep Learning**. Introduced a novel Deep Convolutional Neural Network detecting 20 key-point facial features for recognition, achieving state-of-the-art results. The system demonstrated real-time performance on a UAV, operating at 19 FPS.

ONGOING PROJECTS

- Test-Time Adaptation of Vision-Language models using LoRA (in collaboration)**. An innovative approach employing LoRA adapters is proposed for conducting test-time adaptation of CLIP, enhancing its capability for zero-shot recognition on novel domains. (ongoing submission to **ECCV 2024**).
- Temporally consistent Video generation using Latent consistency models**. A new computation efficient method for video generation using consistency models for generating temporally coherent videos.
- A Foundational Model for Agriculture** (in collaboration). A specialized foundational model for effectively classifying diverse fine-grained agricultural categories.

PUBLICATIONS

- Hanan Gani**, Muzammal Naseer, Salman Khan and Fahad Khan. “*MedContext: Learning Contextual Cues for Efficient Volumetric Medical Segmentation*”. Under review at **CVPR 2024**.
- Hanan Gani**, Shariq Farooq, Muzammal Naseer, Salman Khan and Peter Wonka. “*LLM Blueprint: Enabling Text-to-Image Generation with Complex and Detailed Prompts*”. Accepted at 12th **International Conference on Learning Representations (ICLR) 2024**.
 paper: [arXiv:2310.10640](https://arxiv.org/abs/2310.10640)  [Code](#)
- Hanan Gani***, Jameel Hassan*, Noor Hussein, Mohammad Uzair Khattak, Muzammal Naseer, Salman Khan and Fahad Khan. “*Align Your Prompts: Test-Time Prompting with Distribution Alignment for Zero-Shot Generalization*”. In proceedings of 37th **Advances in Neural Information Processing Systems (NeurIPS) 2023**.
 paper: [arXiv:2311.01459](https://arxiv.org/abs/2311.01459)  [Code](#)
- Hanan Gani**, Muzammal Naseer, Mohammad Yaqub. “*How To Train Vision Transformer On Small-scale Datasets?*”. In proceedings of 33rd **British Machine Vision Conference (BMVC)**, UK, 2022.
 paper: [arXiv:2210.07240](https://arxiv.org/abs/2210.07240)  [Code](#)
- S. Kumar, A. Majeedi, A. Dogra, **H. Gani**, R. M. Vishwanath and S N Omkar. “*Disguised Facial Recognition using Neural Networks*”. **IEEE 3rd International Conference on Signal and Image Processing (ICSIP)**, Shenzhen, China, 2018, pp. 28-32. doi: 10.1109/SIPROCESS.2018.8600440
- Saumya Kumar, Abrar Majeedi, **Hanan Gani**, Abhinandan Dogra, Ravi M. Vishwanath and S N Omkar. “*A Supervised learning Methodology for Real time Disguised Facial Recognition in Wild*”. Accepted to **2018 ACM International Conference on Robotics and Computer Vision (ICRCV)**.
 paper: [arXiv:1809.02875](https://arxiv.org/abs/1809.02875)  [Code](#)

PATENTS

- Hanan Gani, Muzammal Naseer, Mohammad Yaqub. “*System and Method of Training Vision Transformer on Small-Scale Datasets*”. USPTO application no.: 18089107. Passed all three stages of assessment. **US Patent** filed (in process).

TECHNICAL AND PROGRAMMING SKILLS

ML and deep learning Libraries & Frameworks: Pytorch, Keras, Tensorflow, OpenCV, Scikit-learn Python programming, Python for Machine learning and Data Science MATLAB, SciLab (Limited proficiency) C Programming, HTML, Databases: {MySQL, NoSql MongoDB}, WebAPI Hosting, C#, Flask.

RELEVANT UNIVERSITY COURSEWORK AND MOOC'S TAKEN

MSc. Credit Courses: Machine Learning (ML-701), Statistical Inference and Causality (ML-703), Foundations of Artificial Intelligence (AI-701), Mathematics (MTH-701), Trustworthy Artificial Intelligence (ML-708)- MSc. Credit Courses
 Undergraduate credit courses: Random Processes (ECE-505), Image Processing (ECE-019E), Mathematics (MTH-101, 201, 306, 403)
 MOOCs with certifications (coursera.org): Build Generative Adversarial Networks; AI for medical diagnosis; Deep Learning 16 weeks specialization; Machine learning 24 weeks specialization; Data Science crash course; programming and data structures

AWARDS, SCHOLARSHIPS, ACHIEVEMENTS AND INVOLVEMENTS

Awarded **NeurIPS 2023** Travel Grant. *September 2023*
 Serving as a Reviewer at CVPR 2024, ICLR 2024, NeurIPS 2023 and ICML 2023. *March 2023 -*
 My work on **Multi-Task Learning in Vision Transformers** got accepted as an **Oral** paper at UAE GSRC 2023. *March 2023*
 Selected as one of the few candidates to participate in the Google India Research Week 2022. *Jan 2022*
 Received Harman Star Excellence award from the Harman International (Global Test Automation) India (Regional) Head for developing two machine learning solutions which are currently helping the Automation teams in India to save a time effort of 2 hours daily *September 2020*
 Merit Based Scholarship granted for undergraduate studies by Ministry of Minority Affairs, India. *August 2016 - April 2018*
 Secured 80th state rank in IIT-JEE Mains 2014 (among top 1% of 1.5 million students across the country). *June 2014*
 Best Outgoing student of the school. *November 2013*

SOCIAL CAUSE AND VOLUNTEERSHIP

- **'Rivero'** - An initiative for Social Change *Baramulla, Kashmir, India*
Co-Founder *Dec 2016 - Present*
Highlights: Rivero is an NGO based in Kashmir which aims at counseling students for various career options and conducting events and workshops for expressing ideas to bring about a social change. Rivero is pretty successful in conducting numerous educational events and workshops and counsel up-to 2000 students till now with majority being underprivileged and conflict affected students of Kashmir.

EXTRACURRICULAR ACTIVITIES & HOBBIES

Active participation in trekking, camps, and sports activities such as cricket, table tennis, football, badminton etc.
 Social Networking and Communication
 Watching sports activities
 Reading technological stuff

REFERENCES

1. **Dr. Salman Khan**, Associate Professor, Mohamed Bin Zayed University of Artificial Intelligence & Australian National University (ANU)
 salman.khan@mbzuai.ac.ae
2. **Dr. Fahad Khan**, Professor and Deputy Chair Computer Vision Department, Mohamed Bin Zayed University of Artificial Intelligence & Linkoping University, Abu Dhabi, UAE
 fahad.khan@mbzuai.ac.ae
3. **Dr. Muzammal Naseer**, Research Scientist, Mohamed Bin Zayed University of Artificial Intelligence, Abu Dhabi, UAE
 muzammal.naseer@mbzuai.ac.ae
4. **Dr. Peter Wonka**, Professor and Associate Director of VCC, King Abdullah University of Artificial Intelligence (KAUST), Saudi Arabia
 pwonka@gmail.com
5. **Dr. Kun Zhang**, Associate Professor, Carnegie Mellon University (CMU) and MBZUAI
 kunz1@cmu.edu
6. **Dr. Mohammad Yaqub**, Associate Professor, Mohamed Bin Zayed University of Artificial Intelligence, Abu Dhabi, UAE
 mohammad.yaqub@mbzuai.ac.ae
7. **Dr. Abubakar Abid**, Machine Learning Lead, Hugging Face Inc, USA
 a12d@stanford.edu