Mohammad Hanan Gani

☎(+971)585362287 | ✉ hanan.ghani@mbzuai.ac.ae | ✉ m.hanan3829@gmail.com | Website | Github | Google Scholar

Research Interests

I am interested in addressing the challenges faced by multi-modal AI systems across discriminative, generative, and perceptual understanding tasks, particularly when confronted with Out-Of-Distribution (OOD) scenarios.

Education

Mohamed Bin Zayed University of Artificial Intelligence (MBZUAI) Abu Dhabi, UAE
Master of Science (M.Sc.) in Machine Learning, GPA: 3.77/4.0 *August. 2022 – May 2024*
(*15th in CS rankings in AI/ML/CV/NLP*)
Advisor: Prof. Salman Khan, *Co-Advisor:* Prof. Fahad Khan
Thesis title: Text-to-Image Diffusion with Complex and Detailed Prompts (link)

National Institute of Technology (NIT) Srinagar, Kashmir, India
Bachelor of Technology (B.Tech), Electronics & Communication Engineering, GPA: 8.561/10 *Aug. 2014 – June 2018*
Advisor: Dr. G. R. Begh, *Co-Advisor:* Dr. Shahid Mehraj
Thesis title: Channel Estimation in Cognitive Radio using Machine Learning

Publications

(* indicates joint first authors, + indicates my role as mentor)

Multi-modal discriminative Learning

- **Hanan Gani**[*], Jameel Hassan[*], Noor Hussein, M. Uzair Khattak, Muzammal Naseer, Fahad Khan and Salman Khan. *"Align Your Prompts: Test-Time Prompting with Distribution Alignment for Zero-Shot Generalization"*. 37th **Advances in Neural Information Processing Systems (NeurIPS) 2023**. Paper Code
- Raza Imam, **Hanan Gani+**, Mohammad Huzaifa and Karthik Nandakumar. *"Test-time Low Rank Adaptation via Confidence Maximization for Zero-Shot Generalization in Vision-Language Models"*. **IEEE/CVF Winter Conference on Applications of Computer Vision (WACV) 2025**. Paper Code
- Umair Nawaz, Awais M., **Hanan Gani+**, M. Naseer, F. Khan, S. Khan and R. Anwer. *"AgriCLIP: Adapting CLIP for Agriculture and Livestock via Domain-Specialized Cross-Model Alignment"*. **COLING 2025** (Accepted).
Paper Code

Multi-modal generative Learning

- **Hanan Gani**, Shariq Farooq, Muzammal Naseer, Salman Khan and Peter Wonka. *"LLM Blueprint: Enabling Text-to-Image Generation with Complex and Detailed Prompts"*. 12th **International Conference on Learning Representations (ICLR) 2024**. Paper Code

Multi-modal perception & understanding

- Shehan Munasinghe[*], **Hanan Gani+**[*], Wenqi Zhu, Jiale Cao, Eric Xing, Fahad Khan and Salman Khan. *"VideoGLaMM: A Large Multimodal Model for Pixel-Level Visual Grounding in Videos"*.
Under review at **CVPR 2025**. Paper Code
- **Hanan Gani***, Rohit Bhardwaj*, Muzammal Naseer, Fahad Khan and Salman Khan. *"VANE-Bench: Video Anomaly Evaluation Benchmark for Conversational LMMs"*. Under review at **NAACL 2025**.
Paper Code

Sample-Efficient Learning and Miscellaneous

- **Hanan Gani**, Muzammal Naseer, Fahad Khan and Salman Khan. *"MedContext: Learning Contextual Cues for Efficient Volumetric Medical Segmentation"*. 27th **International Conference on Medical Image Computing and Computer Assisted Intervention Society (MICCAI) 2024**. Paper Code
- **Hanan Gani**, Muzammal Naseer, Mohammad Yaqub. *"How To Train Vision Transformer On Small-scale Datasets?"*. 33rd **British Machine Vision Conference (BMVC) 2022**. Paper Code

- **Hanan Gani**, Nada Saadi, Noor Hussein, Karthik Nandakumar. *"Multi-Attribute Vision Transformers are Efficient and Robust Learners"*. Accepted at **IEEE International Conference on Image Processing (ICIP) 2024.**
  Paper Code
- Saumya Kumaar, 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 at **International Conference on Robotics and Computer Vision (ICRCV) 2018**.
  Paper code

Patents

Hanan Gani, Muzammal Naseer, Mohammad Yaqub. *"System and Method of Training Vision Transformer on Small-Scale Datasets"*. **US Patent**. Pub. No. US 2024/0212330 A1. USPTO application no.: 18089107

Research Experience

King Abdullah University of Science and Technology (KAUST) Saudi Arabia
Visiting Sudent *June 2023 – October 2023*
Host Advisor: Prof. Peter Wonka, Full Professor, Computer Science Department

- Worked on **LLM Blueprint**, the first approach enabling diffusion models to generate detailed, complex scenes from lengthy and intricate textual prompts, significantly advancing text-to-image generation capabilities (*ICLR 2024*).
- On the 2-AFC task, LLM Blueprint was picked by an average of 69.3% of participants as coherent image generator.

Mohamed Bin Zayed University of Artificial Intelligence (MBZUAI) Abu Dhabi, UAE
Research Assistant *Oct 2021 – Aug. 2022*
Advisors: Prof. Mohammad Yaqub, Associate Professor, Computer Vision Department
Prof. Muzammal Naseer, Assistant Professor, Khalifa University (Prev. Research Scientist at MBZUAI)

- Developed a self-supervised scheme to learn weights from low-resolution views on small datasets, enabling Vision Transformers to be trained from scratch without large-scale pre-training (*BMVC 2022 and US Patent*).
- Proposed approached can be used in a plug-and-play manner without any changes to architecture or loss function.

Fatima Fellowship – Predoc Program USA (remote)
Part-time fellow *March 2021 – Dec. 2021*
Advisor: Dr. Abubakr Abid, ML Lead, Hugging Face Inc

- Developed a multi-task approach optimizing Vision Transformers to efficiently handle multiple tasks simultaneously within a constrained computational budget without compromising performance. (*IEEE ICIP'24, UAE GSRC'23*)

Indian Institute of Science (IISc) Bangalore, India
Research Intern *Dec. 2017 – March 2018*
Host Lab: Computational Intelligence & UAV Lab, Aerospace Engineering Department

- Worked with a team in developing a lightweight Disguised Facial Recognition System utilizing 20 unique facial keypoints for recognition, capable of operating in real-time at 19 FPS (*IEEE ICSIP'18, ICRCV'18*).

Work Experience

Mohamed Bin Zayed University of Artificial Intelligence (MBZUAI) Abu Dhabi, UAE
Research Associate II *June. 2024 – Present*
Advisor / Line Manager: Prof. Salman Khan

- Working on developing an image super-resolution model for low-quality satellite sensory imagery for UAE region to determine cloud coverage optimal cloud seeding, reducing operational costs (collaboration with UAE Govt.).
- Developed a spatiotemporal Video Grounding model with an MS student capable of fine-grained grounded conversation generation, visual grounding, and referring object segmentation (*CVPR'25* under review).
- Developed a new benchmark for evaluating Video LMMs on inconsistencies in AI generated videos (*NAACL'25*)

Samsung R&D Bangalore, India
Machine Learning Engineer *Oct. 2018 – Sept. 2021*

- Developed ***Screen Reliability system*** for real-time detection of anomalies in continuous video streams on HMI screens using deep learning models such as Auto-encoders and GANs (used in production).

- Developed ***Test Case Recommender*** which uses transformer based language model to map user query with the relevant test cases to fix automation issues such as software failures or system crashes. (used in production)
- Developed ***Log Failure Classifier***, an ML model which utilizes error logs to distinguish between software and hardware failures. (saves 2 hours per day to software team)
- Devised a tool called ***Similar Issue Recommender***, which uses language model to recommend fixes to software issues based on past fix history. (used in production)

Teaching and Academic Services

Teaching Assistant | *MBZ University of Artificial Intelligence (MBZUAI), Abu Dhabi*
- Advanced Topics in Vision and Language (CV806) – Spring 2024 with Prof. Ivan Laptev
- Deep Learning (AI702) – Spring 2024 with Prof. Harris Khan
- Probabilistic and Statistical Inference (ML703) – Fall 2023 with Prof. Kun Zhang
- Machine Learning (ML701) – Fall 2023 with Prof. Samuel Horvath

Tutor & Lab Instructor | *UGRIP program, MBZUAI*
- Tutor for Foundations of AI course and lab instructor for UGRIP internship program at MBZUAI

Conference Reviewing | *Reviewer*
- ICLR'25, 24 & 23, AAAI'24, WACV'24, ECCV'24, NeurIPS'24 & 23, CVPR'24, ICML'24, ICCV'23

Conference Volunteer | *IEEE International Conference on Image Processing (ICIP) 2024*
- Serving as a program volunteer for IEEE ICIP conference in Abu Dhabi, UAE

Invited Talk | *University of Jordon*
- Presented a talk on Generative AI to undergrad CS students at the University of Jordon

Pilot Mentorship Program, MBZUAI | Mentor
- As an alumnus, I am mentoring 3 master's students from the 2025 graduating batch of MBZUAI, providing them mentorship regarding their final project thesis and potential research publications.

Honors and Awards

- Honored by MBZUAI for my research contributions during masters degree (link)
- Awarded **ICLR 2024** Travel Grant
- Awarded **NeurIPS 2023** Travel Grant
- MBZUAI graduate studies scholarship holder
- Selected as one of the few candidates from India to participate in the Google India Research Week 2022
- Received Samsung-Harman Star Excellence award from India Regional Head
- Merit Based Scholarship granted for undergraduate studies by Ministry of Minority Affairs India.

Technical Skills

Languages: Python, C, C#, MATLAB, SQL, HTML/CSS
Frameworks: PyTorch, TensorFlow, Keras, Scikit-Learn, OpenCV, HuggingFace, Flask
Developer Tools: Git, Docker, Google Cloud Platform, VS Code, Visual Studio, PyCharm, Linux

Extracurricular and Social Activities

Graduate Student Council (GSC) | Machine Learning Coordinator & Member of GSC at MBZUAI
- Act as a bridge between the students and department and discuss any student-related issues with the head of the ML department to ensure issues are addressed promptly.
- Organize general student-centric initiatives and programs with the administration as a part of GSC team

Rivero | Co-Founder (undergrad initiative)
- Rivero is a Kashmir-based NGO dedicated to guiding students in exploring various career paths. Through this initiative, we organized numerous educational workshops and events, providing counseling to nearly 2,000 students, primarily those from underprivileged and conflict-affected backgrounds.

References

Prof. Salman Khan | ✉
Prof. Muzammal Naseer | ✉
Prof. Kun Zhang | ✉
Dr. Abubakr Abid | ✉
Prof. Fahad Khan | ✉
Prof. Peter Wonka | ✉
Prof. Karthik Nandakumar | ✉