

# Hanan Gani

La Jolla, CA · ✉ hgani@ucsd.edu · ☎ (+1)-619-953-8366 · 🌐 Homepage · 🎓 GoogleScholar ·  
📁 github.com/hananshafi · in linkedin.com/hanangani

## RESEARCH INTERESTS

---

Intelligent Multi-modal Systems, World Models, Spatial Reasoning, Embodied AI

## EDUCATION

---

### University of California, San Diego (UCSD)

La Jolla, USA

Powell Fellow, Ph.D. in Computer Science

Sept. 2025 –

Advisor: **Manmohan Chandraker** · Fellowships: Jacobs Engineering Fellowship

Specific focus areas: Multi-modality–World Models, Embodiment

### Mohamed Bin Zayed University of Artificial Intelligence (MBZUAI)

Abu Dhabi, UAE

Master of Science (M.Sc.) in Machine Learning, GPA: 3.77/4.0

Aug. 2022 – May 2024

Advisor: **Salman Khan**, Co-Advisor: **Fahad Khan**

Thesis: Text-to-Image Diffusion with Complex and Detailed Prompts

### 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: **G. R. Begh**, Co-Advisor: **Shahid Mehraj**

Thesis: Channel Estimation in Cognitive Radio using Machine Learning

## PUBLICATIONS

---

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

### Multi-modal Grounding, Reasoning & Understanding

- [Under Review] Ghazi S. Ahmed, Ahmed Heakl, **Hanan Gani**+, Abdelrahman Shaker, Zhiqiang Shen, Fahad Khan and Salman Khan. “VideoMolmo: Spatio-Temporal Grounding meets Pointing”. [Paper] [Code]
- [Under Review] Tajamul Ashraf\*, Amal Saqib\*, **Hanan Gani**, Muhra AlMahri, Yuhao Li, Noor Ahsan, Umair Nawaz, Jean Lahoud, Hisham Cholakkal, Mubarak Shah, Philip Torr, Fahad Khan, Rao M. Anwer, Salman Khan. “Agent-X: Evaluating Deep Multimodal Reasoning in Vision-Centric Agentic Tasks”. [Paper] [Code]
- [ICCV 2025] Sanjoy Chowdhury\*, **Hanan Gani**\*, Nishit Anand, S. Nag, Ruohan Gao, Mohamed Elhoseiny, Salman Khan, Dinesh Manocha. “AURELIA: Test-time Reasoning Distillation in Audio-Visual LLMs”. [Paper] [Code]
- [CVPR 2025] Shehan Munasinghe\*, **Hanan Gani**+, Wenqi Zhu, Jiale Cao, Eric Xing, Fahad Khan, Salman Khan. “VideoGLaMM: A Large Multimodal Model for Pixel-Level Visual Grounding in Videos”. [Paper] [Code]
- [NAACL 2025] **Hanan Gani**\*, Rohit Bhardwaj\*, Muzammal Naseer, Fahad Khan and Salman Khan. “VANE-Bench: Video Anomaly Evaluation Benchmark for Conversational LLMs”. [Paper] [Code]

### Multi-modal Discriminative Learning

- [NeurIPS 2023] Jameel Hassan\*, **Hanan Gani**\*, N. Hussein, Uzair Khattak, M. Naseer, Fahad Khan, Salman Khan. “Align Your Prompts: Test-Time Prompting with Distribution Alignment for Zero-Shot Generalization”. [Paper] [Code]
- [WACV 2025] 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”. [Paper] [Code]
- [COLING 2025] U. Nawaz, Awais M., **Hanan Gani**+, M. Naseer, F. Khan, S. Khan and R. Anwer. “AgriCLIP: Adapting CLIP for Agriculture & Livestock via Domain-Specialized Cross-Model Alignment”. [Paper] [Code]

### Multi-modal Generative Learning

- [ICLR 2024] **Hanan Gani**, Shariq Farooq, Muzammal Naseer, Salman Khan and Peter Wonka. “LLM Blueprint: Enabling Text-to-Image Generation with Complex and Detailed Prompts”. [Paper] [Code]

## Sample-Efficient Learning and Miscellaneous

- [MICCAI 2024] **Hanan Gani**, Muzammal Naseer, Fahad Khan and Salman Khan. “*MedContext: Learning Contextual Cues for Efficient Volumetric Medical Segmentation*”. [Paper] [Code]
- [BMVC 2022] **Hanan Gani**, Muzammal Naseer, Mohammad Yaqub. “*How To Train Vision Transformer On Small-scale Datasets?*”. [Paper] [Code]
- [IEEE ICIP 2024] **Hanan Gani**, Nada Saadi, Noor Hussein, Karthik Nandakumar. “*Multi-Attribute Vision Transformers are Efficient and Robust Learners*”. [Paper] [Code]
- [IEEE ICRCV 2018] S. Kumaar, A. Majeedi, **H. Gani**, A. Dogra, Ravi M. Vishwanath, S.N. Omkar. “*A Supervised learning Methodology for Real time Disguised Facial Recognition in Wild*”. [Paper] [Code]

## PATENTS

**Hanan Gani**, Muzammal Naseer, Mohammad Yaqub. “*System and Method of Training Vision Transformer on Small-Scale Datasets*”. [US Patent].

## RESEARCH AND INDUSTRY EXPERIENCE

**MBZ University of Artificial Intelligence (MBZUAI) – Research Associate II** Abu Dhabi, UAE  
Advisor / Line Manager: **Salman Khan** June. 2024 – Sept. 2025

- Co-designed **VideoMolmo** for fine-grained point-level grounding; **+5.4 pp** improvement compared to SOTA on proposed VPoS-Bench; open-source code (*Under review*).
- Co-developed **VideoGLaMM** for grounded conversation, visual grounding, & referring segmentation; **+33.74 pp mIoU** improvement over SOTA for GCG task on proposed benchmark; open-source code (*CVPR’25*).
- Developed super-resolution model for satellite sensory imagery for cloud-seeding ops (UAE Govt. collaboration).
- Developed **VANE-Bench** for LMM inconsistencies in AI videos (*NAACL’25*; released code & eval suite).

**King Abdullah University of Science & Tech. (KAUST) – Visiting Student** Saudi Arabia  
Host Advisor: **Peter Wonka**, Full Professor, CS department June 2023 – Oct. 2023

- Developed **LLM Blueprint** for long, complex T2I diffusion; **69.3% 2-AFC** user preference vs. baselines; open-source release (*ICLR’24*).

**MBZ University of Artificial Intelligence (MBZUAI) – Research Assistant** Abu Dhabi, UAE  
Advisors: **Mohammad Yaqub, Muzammal Naseer** Oct. 2021 – Aug. 2022

- Developed a self-supervised scheme enabling ViTs to train from scratch on small datasets; plug-and-play design; **>90% top-1 acc.** with only 25% train data on CIFAR10 (*BMVC’22, US Patent*).

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

- Developed a multi-task ViT to efficiently and robustly handle multiple tasks simultaneously within a constrained computational budget; **+59.1%** gain in robust acc. w.r.t baseline under Patch-Fool attack (*IEEE ICIP’24*).

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

- Developed Autoencoder/GAN-based online **anomaly detection** for HMI video streams; **90%** fewer false alarms.
- Developed transformer-based **Recommender Systems** for software fixes (used in production).
- Developed **Log Classifier** to distinguish software & hardware failures; saved **~2 hrs/day** to software team
- Co-developed **Multi-agent UI testing** based on icon detection, screen classifier, OCR (used in production).

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

- Co-developed lightweight disguised-face recognition model; **20 keypoints, 19 FPS** (IEEE ICSIP’18, ICRCV’18).

## TECHNICAL SKILLS

**Languages:** Python, C, C#, MATLAB, SQL

**Frameworks:** PyTorch, HuggingFace, TensorFlow, Keras, Scikit-Learn, OpenCV

**ML Infrastructure & Tooling:** Distributed training, DeepSpeed, Accelerate, PyTorch Lightning, W&B, TensorBoard

**Dev Tools:** Linux, Git, Google Cloud Platform, VS Code, PyCharm, Linux, Flask

## TEACHING AND ACADEMIC SERVICES

---

### Teaching | MBZUAI

- TA for **CV806** (Advanced Topics in Vision & Language, Spring'24), **AI702** (Deep Learning, Spring'24), **ML703** (Probabilistic & Statistical Inference, Fall'23), **ML701** (Machine Learning, Fall'23).
- Tutor & Lab Instructor for **UGRIP program** at MBZUAI (Summer'24).

### Conference Services

- **Regular Reviewer:** NeurIPS, ICLR, CVPR, ICCV, ECCV, ICML, AAAI
- **Program Volunteer:** IEEE ICIP'24 (Abu Dhabi, UAE)

### Invited Talks

- **Abu Dhabi University**, May'25: Delivered a talk on “Latest Trends in AI” at CS department.
- **University of Jordan**, June'24: Presented a talk on “Generative AI” to undergrad CS students.

### Pilot Mentorship Program, MBZUAI | Mentor, 2024-2025

- Mentored 3 MS students in thesis and publication guidance via MBZUAI Pilot Mentorship Program.

## HONORS AND AWARDS

---

- Recipient of prestigious **Powell Fellowship** and **Jacobs Engineering Fellowship** at UC San Diego.
- 2<sup>nd</sup> place in Research track at UC Berkeley's **AgentX Competition**.
- Awarded **CVPR 2025** Travel Award
- Honorable mention 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
- Received Samsung-Harman Star Excellence award from India Regional Head
- Merit Based Scholarship granted for undergraduate studies by Ministry of Minority Affairs India.

## EXTRACURRICULAR AND SOCIAL ACTIVITIES

---

**Graduate Student Council** | Machine Learning Coordinator; liaison with department leadership; student initiatives.

**Rivero** | Co-Founder; NGO guiding students in Kashmir; workshops and counseling for ~2,000+ underprivileged students