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
  Main Supervisor: Dr. Salman Khan, Associate Professor (✉ Salman.Khan@mbzuai.ac.ae)
  Secondary Supervisor: Dr. Fahad Khan, Associate 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; Zero-Shot Prompt learning; Multi-modal learning ; 3D Segmentation
  Academic work: Teaching Assistant Aug 2023 - Present
  1. *Machine Learning (ML701)*: Course Instructor - Dr. Samuel Horvath
  Duties: Instruct lab sessions for the course; grade assignments, exams and quizzes; mentor students for the course projects.
  2. *Probabilistic and Statistical Inference (ML703)*: Course Instructor - Dr. Kun Zhang
  Duties: Instruct lab sessions for the course; provide feedback on student progress; grade assignments and exams; project guidance.

- **National Institute of Technology (NIT)** *Srinagar, India*
  Bachelor of Technology (B.Tech), Electronics and Communication Engineering 2014-2018
  Overall GPA: 8.561/10 (Among top 5 of the class)

- **Saint Joseph's Higher Secondary School** *Baramulla, J&K (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 AND RESEARCH EXPERIENCE

- **King Abdullah University of Science and Technology (KAUST)** *Thuwal city, Saudi Arabia*
  Visiting Student June 2023 - August 2023
  Advisor: Dr. Peter Wonka, Professor and Associate Director of Visual Computing Center (VCC), Computer Science Department, (peter.wonka@kaust.edu.sa)
  Highlights of Research:
  □ **Text-to-Image generation from longer 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 text description.

- **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)*
  Fatima Fellow April 2021 - Dec 2021
  Mentor: Dr. Abubakar Abid, Machine Learning Lead at Hugging Face Inc (USA), Founder at Gradio Inc, PhD at Stanford University (a12d@stanford.edu)
  Highlights of Research: **Multi-Task learning (MTL)** is a challenging research area in deep learning. Under Fatima Fellowship, I worked with Dr. Abubakar Abid in exploring new realms of doing smart Multi-Task learning (MTL) in order to make AI algorithms capable of doing multiple tasks at a time with limited computational cost. Our approach is based on exploiting the class-token and self-attention mechanism of Vision Transformers (ViT's) in order to train multiple tasks through a single ViT, more efficiently and with limited computational budget. Selected as **Oral** paper at UAE Graduate Student Research Conference (GSRC) (Project demo code: https://github.com/hananshafi/MTL-ViT).

- **Harman International - Connected Car R&D (Samsung)** *Bengaluru, India*
  Machine Learning Research Engineer Oct 2018 - September 2021
  Subdivision: Harman Connected Car R&D
  Subgroup: Global Test Automation (GTA) - Machine Learning R&D Team
  Projects and Research work:
  □ **Screen Reliability - detecting anomalies on HMI screens**: The project is based on detecting anomalies in a continuous video stream (on HMI screen) using deep learning based anomaly detection. Our approach is based on using a convolutional Auto-encoder network and conditional Generative Adversarial Network inspired from the 'GANomaly' paper. (currently being used in production at Harman facilities)
  □ **Test Case Recommender: Mapping contextually similar texts together** : The project has been integrated to fix the automation issues faced by the company on daily basis, which can be fixed by running the relevant test cases pertaining to the issues. In case of a software

run failure / crash, our SBERT based machine learning model maps the prompt from the failure logs to the relevant test cases, thereby fixing the failure scenarios (currently being used in production at Harman facilities)

☐ *Log Failure Categorization*: This project is based on extracting useful information regarding the cause of failure from plethora of error logs generated from various test executions. A machine learning pipeline is developed which classifies failure text logs into two categories depending on whether the failure was from software side or hardware side. (currently being used in production at Harman facilities)

☐ *Hybrid Icon Detection Model*: Developed a Hybrid Deep Learning Icon Detection Model that can detect various icons on HMI screen images. The model is robust to any kind of changes like background variations, screen or icon scaling, resolution changes and icon shape changes. The hybrid implementation combines the results from deep learning model and statistical machine learning model, and gives the final bounding box outputs. (currently being used in production at Harman facilities)

☐ *Word to Vec similar Issue Recommender*. This project is based on recommending similar issue fixes in the past given a new issue. A software issue in the form of sentence prompt is fed as input and a Word2Vec model (trained on large corpus of data) converts this information into an embedding vector which is matched with the embedding vectors of previous records in the database. On the basis of cosine similarity score, the model recommends similar types of software issues that had been 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)

- **Indian Institute of Science (IISc)**   *Bengaluru, India*
  Deep Learning and Computer Vision Researcher   Full time: Dec 2017 - Feb 2018 , Part time: March 2018 - June 2018
  Lab: Computational Intelligence & UAV Lab, Aerospace Engineering Department, IISc
  Highlights of Research: Carried out extensive research in Deep Learning and Computer Vision and worked on Project **Disguised Facial Recognition using Deep Learning**. This research project presents a novel approach for disguised facial recognition using a novel Deep Convolutional Neural Network which detects 20 essential key-point features on face. These 20 key-point features are then utilized by an artificial neural network for recognition task. The performance achieved state of the art results. The system is also tested in real time on a UAV, working at 19 FPS, thus almost performing in real time.

- **National Institute of Technology (NIT)**   *Srinagar, India*
  Student Researcher   October 2017 - June 2018
  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)
  Project work: ☐ *Machine learning based channel estimation*: Developed an efficient Machine Learning based method to estimate Channel Parameter 'H' in Wireless Communication System. 'H' parameter presents sum total of all the factors influencing the input signal when it travels from source to receiver and is represented as Gaussian Noise. The motive is to get the original sent input. To accomplish this, I exploited the Machine Learning technique of Least Squares Estimation to estimate 'H' parameter. (full Proficiency).
  ☐ Developed a **Real time Emotion Recognition System** that recognizes five types of emotions from facial expressions: sad, happy, angry, surprise and neutral. A standard (sequential) and inception style CNN architectures were deployed and separately trained on FER benchmark dataset. The test set results were used to compare the performance of the two architectures.

PUBLICATIONS

- **Hanan Gani**[*], Shariq Farooq, Muzammal Naseer, Salman Khan and Peter Wonka. *"LLM Blueprint: Enabling Text-to-Image Generation with Elaborate and Detailed Prompts"*. Under review at **ICLR 2024**.

- **Hanan Gani**[*], Muzammal Naseer, Salman Khan and Fahad Khan. *"3D-MSR: Masked Segmentation Reconstruction for Label Efficient Volumetric Medical Segmentation"*. Under review at **AAAI 2024**.

- **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"*. 37th **Advances in Neural Information Processing Systems (NeurIPS) 2023**.

- **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. **arXiv:2210.07240 [cs.CV]**

- S. Kumaar, 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/SIPRO-CESS.2018.8600440

- 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"*. Published on **arXiv:1809.02875[cs.CV]**. Accepted to **2018 ACM International Conference on Robotics and Computer Vision (ICRCV)**, Nov 17-18, Thailand.

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).

RESEARCH INTERESTS

☐ Vision-Language Models ☐ Self-supervised learning ☐ Open-World Semantic Segmentation ☐ Multi-Task Learning ☐ Data-Efficient Vision Transformers ☐ Causal Inference in Machine Learning ☐ 3D Computer Vision ☐ 3D Medical imaging

TECHNICAL AND PROGRAMMING SKILLS

☐ ML and deep learning Libraries & Frameworks: Keras, Tensorflow, Pytorch, 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

☐ Serving as a Reviewer at 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, Abu Dhabi, UAE
   ✉ salman.khan@mbzuai.ac.ae

2. **Dr. Fahad Khan**, Professor and Deputy Chair Computer Vision Department, Mohamed Bin Zayed University of Artificial Intelligence, 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. Abubakar Abid**, Machine Learning Lead, Hugging Face Inc, USA
   ✉ a12d@stanford.edu

5. **Dr. Mohammad Yaqub**, Associate Professor, Mohamed Bin Zayed University of Artificial Intelligence, Abu Dhabi, UAE
   ✉ mohammad.yaqub@mbzuai.ac.ae

6. **Dr. Shahid Mehraj**, Assistant Professor, National Institute of Technology (NIT) Srinagar
   ✉ shahidshah@nitsri.net

7. **Dr. G. R. Begh**, Associate Professor, National Institute of Technology (NIT) Srinagar
   ✉ grbegh@nitsri.ac.in