

Muhammad Akhtar Munir

Research Scientist | Multimodal & Vision-Language Models, Foundation Models,
Reliable ML, Agentic Methods

its.akhtar90@gmail.com | akhtar.munir@mbzuai.ac.ae
[Google Scholar](#) | [LinkedIn](#) | [Website](#)

Summary

I am a Research Scientist in Computer Vision at Mohamed bin Zayed University of Artificial Intelligence (MBZUAI), Abu Dhabi. My work spans multimodal learning, vision-language models, foundation models, model calibration, robust evaluation, and agentic methods for complex real-world data.

My research includes foundation models for multimodal data, vision-language model evaluation and adaptation, calibrated prediction under domain shift, benchmark design, and tool-augmented reasoning. I have mainly used earth observation, geospatial data, climate, and environmental monitoring as application domains, alongside work on VLM calibration and domain-shift robustness. My publications include CVPR, ICCV, NeurIPS, ICLR, ECCV, and TPAMI, with CVPR/ICCV highlight papers and Best Paper recognition at ICML TerraBytes and the CVPR MONTI Workshop.

Experience

Research Scientist

Mar 2026 – Present

Mohamed bin Zayed University of Artificial Intelligence (MBZUAI), Abu Dhabi, UAE

Department of Computer Vision

- Lead research on multimodal foundation models, vision-language models, reliable computer vision, and tool-augmented/agentic methods.
- Drive problem formulation, experimental design, benchmark development, evaluation protocols, ablation studies, and reproducible research workflows.
- Develop methods combining visual perception, language, structured reasoning, and external tools; mentor research engineers and PhD students across collaborative projects.

Postdoctoral Associate II

Jul 2025 – Feb 2026

Mohamed bin Zayed University of Artificial Intelligence (MBZUAI), Abu Dhabi, UAE

Department of Computer Vision

- Worked on foundation models, vision-language models, multimodal learning, and benchmark design for visual and geospatial reasoning.
- Contributed to evaluation frameworks and research prototypes for multimodal, foundation-model, and tool-augmented methods.

Postdoctoral Associate

Jul 2023 – Jun 2025

Mohamed bin Zayed University of Artificial Intelligence (MBZUAI), Abu Dhabi, UAE

Department of Computer Vision

- Conducted research in multimodal learning, reliable machine learning, Earth observation, climate-focused modeling, and evaluation under real-world distribution shifts.
- Developed weather and air-quality forecasting methods with emphasis on regional adaptation, reliability, and environmental monitoring.

Research Assistant/Intern

Aug 2022 – Jun 2023

Mohamed bin Zayed University of Artificial Intelligence (MBZUAI), Abu Dhabi, UAE
Department of Computer Vision

- Conducted research in object detection, model calibration, domain shift, and generalization.

Researcher

Jan 2022 – Jul 2022

Information Technology University (ITU) of Punjab, Lahore, Pakistan
Under Google Research Scholar Award

- Conducted research in domain generalization and poverty estimation from imagery.

PhD Fellow

Sep 2017 – Aug 2021

Information Technology University (ITU) of Punjab, Lahore, Pakistan
Intelligent Machines Lab

- Conducted doctoral research on domain adaptation, object detection, model calibration, and reliable visual recognition.

Lab Engineer

Nov 2013 – Aug 2017

School of Electrical Engineering and Computer Science (SEECS), NUST, Islamabad, Pakistan
Department of Computing

- Supported teaching, lab development, and practical instruction in programming, computer vision, and artificial intelligence.

Publications

1. Muhammad Sohail Danish, **Muhammad Akhtar Munir**, Syed Roshan Ali Shah, Muhammad Haris Khan, Rao Muhammad Anwer, Jorma Laaksonen, Fahad Shahbaz Khan, Salman Khan. *TerraFM: A Scalable Foundation Model for Unified Multisensor Earth Observation*. International Conference on Learning Representations (ICLR 2026).
2. Akashah Shabbir*, **Muhammad Akhtar Munir***, Akshay Dudhane*, Muhammad Umer Sheikh, Muhammad Haris Khan, Paolo Fraccaro, Juan Bernabe Moreno, Fahad Shahbaz Khan, Salman Khan. *ThinkGeo: Evaluating Tool-Augmented Agents for Remote Sensing Tasks*. MONTI Workshop, IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2026). **[Best Paper]**
***Equally Contributing First Authors.**
3. Akashah Shabbir, Muhammad Umer Sheikh, **Muhammad Akhtar Munir**, Hiyam Debary, Mustansar Fiaz, Muhammad Zaigham Zaheer, Paolo Fraccaro, Fahad Shahbaz Khan, Muhammad Haris Khan, Xiao Xiang Zhu, Salman Khan. *OpenEarthAgent: A Unified Framework for Tool-Augmented Geospatial Agents*. European Conference on Computer Vision (ECCV 2026).
4. Ashshak Sharifdeen, Fahad Shamshad, **Muhammad Akhtar Munir**, Abhishek Basu, Mohamed Insaf Ismithdeen, Jeyapriyan Jeyamohan, Chathurika Sewwandi Silva, Karthik Nandakumar, Muhammad Haris Khan. *Towards Calibrating Prompt Tuning of Vision-Language Models*. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2026).
5. Yohan Abeysinghe, **Muhammad Akhtar Munir**, Sanoojan Baliah, Ron Sarafian, Fahad Shahbaz Khan, Yinon Rudich, Salman Khan. *Synergistic Neural Forecasting of Air Pollution with Stochastic Sampling*. Nature Portfolio: npj Clean Air (2026).

6. Muhammad Sohail Danish*, **Muhammad Akhtar Munir***, Syed Roshaan Ali Shah, Kartik Kuckreja, Fahad Shahbaz Khan, Paolo Fraccaro, Alexandre Lacoste, Salman Khan. *GEOBench-VLM: Benchmarking Vision-Language Models for Geospatial Tasks*. IEEE/CVF International Conference on Computer Vision (ICCV 2025). **[Highlight Paper]**
***Equally Contributing First Authors.**
7. Ashshak Sharifdeen, **Muhammad Akhtar Munir**, Sanoojan Baliah, Salman Khan, Muhammad Haris Khan. *O-TPT: Orthogonality Constraints for Calibrating Test-time Prompt Tuning in Vision-Language Models*. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2025). **[Highlight Paper]**
8. Vishal Nedungadi, **Muhammad Akhtar Munir**, Marc Rußwurm, Ron Sarafian, Ioannis N. Athanasiadis, Yinon Rudich, Fahad Shahbaz Khan, Salman Khan. *AirCast: Improving Air Pollution Forecasting Through Multi-Variable Data Alignment*. TerraBytes Workshop, International Conference on Machine Learning (ICML 2025). **[Best Paper]**
9. Sagar Soni, Akshay Dudhane, Hiyam Debary, Mustansar Fiaz, **Muhammad Akhtar Munir**, Muhammad Sohail Danish, Paolo Fraccaro, Campbell Watson, Levente Klein, Fahad Khan, Salman Khan. *EarthDial: Turning Multi-sensory Earth Observations to Interactive Dialogues*. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2025).
10. **Muhammad Akhtar Munir**, Fahad Khan, Salman Khan. *Efficient Localized Adaptation of Neural Weather Forecasting: A Case Study in the MENA Region*. Workshop on Tackling Climate Change with Machine Learning, NeurIPS 2024.
11. Sohail Danish, Muhammad Haris Khan, **Muhammad Akhtar Munir**, M. Saquib Sarfraz, Mohsen Ali. *Improving Single Domain-Generalized Object Detection: A Focus on Diversification and Alignment*. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2024).
12. **Muhammad Akhtar Munir**, Salman Khan, Muhammad Haris Khan, Mohsen Ali, Fahad Khan. *Cal-DETR: Calibrated Detection Transformer*. Conference on Neural Information Processing Systems (NeurIPS 2023).
13. **Muhammad Akhtar Munir**, Muhammad Haris Khan, M. Saquib Sarfraz, Mohsen Ali. *Domain Adaptive Object Detection via Balancing between Self-Training and Adversarial Learning*. IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI 2023).
14. **Muhammad Akhtar Munir**, Muhammad Haris Khan, Salman Khan, Fahad Khan. *Bridging Precision and Confidence: A Train-Time Loss for Calibrating Object Detection*. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2023).
15. **Muhammad Akhtar Munir**, Muhammad Haris Khan, M. Saquib Sarfraz, Mohsen Ali. *Towards Improving Calibration in Object Detection Under Domain Shift*. Conference on Neural Information Processing Systems (NeurIPS 2022).
16. **Muhammad Akhtar Munir**, Muhammad Haris Khan, M. Saquib Sarfraz, Mohsen Ali. *SSAL: Synergizing between Self-Training and Adversarial Learning for Domain Adaptive Object Detection*. Conference on Neural Information Processing Systems (NeurIPS 2021).

Preprints

1. **Muhammad Akhtar Munir**, Muhammad Umer Sheikh, Akashah Shabbir, Muhammad Haris Khan, Fahad Khan, Xiao Xiang Zhu, Begüm Demir, Salman Khan. *Agentic AI for*

Remote Sensing: Technical Challenges and Research Directions. arXiv 2026.

Patents

- **A Train-time Loss in a System and Method for Calibrating Object Detection.**
Inventors: [Muhammad Akhtar Munir](#), Muhammad Haris Khan, Salman Khan, Fahad Shahbaz Khan.
Patent Office: US. Publication date: Feb 20, 2025. *US20250061697A1*.

Mentorship and Communication

- Mentored PhD students, research engineers, research assistants, and junior collaborators on projects in foundation models, vision-language models, multimodal learning, model calibration, agentic methods, benchmark development, climate modeling, and reproducible experimentation.
- Contributed to project formulation, experimental design, paper writing, revision strategy, research planning, and proposal development across collaborative research projects.
- Supported technical communication and curriculum activities, including labs and guest lectures on lifelong learning, vision-language models, model evaluation, and agentic vision methods at MBZUAI and the University of Central Florida.
- Led undergraduate laboratory sessions and tutorials in programming, computer vision, and artificial intelligence at NUST from 2013 to 2017.

Education

Doctor of Philosophy in Computer Science Sep 2017 – Dec 2024
Information Technology University, Lahore, Pakistan
Thesis: Domain Adaptation for Object Detection
Degree awarded in 2024; requirements completed in 2023.

Master of Science in Computer Science Feb 2015 – Jun 2017
COMSATS Institute of Information Technology, Islamabad, Pakistan
Thesis: Ischemic Stroke Lesion Segmentation in MR Sequences

Bachelor of Science in Computer Engineering Jan 2009 – Jan 2013
COMSATS Institute of Information Technology, Islamabad, Pakistan

Programming Languages & Tools

- **Programming and Deep Learning:** Python, PyTorch
- **Tools and Platforms:** Linux, ArcGIS, VS Code, Git, MS Office, L^AT_EX
- **Research Areas:** Multimodal Learning, Vision-Language Models, Foundation Models, Agentic Methods, Model Calibration, Robust and Reliable Machine Learning, Benchmarking and Evaluation, Earth Observation and Climate Applications

Service, Honors, and Recognition

- **Peer Review Service:** Reviewer for CVPR (2023–2026), ICCV (2023–2025), NeurIPS (2023–2026), ICLR (2024–2026), ECCV (2026), EMNLP (2026), TPAMI, and IJCV.
- **Awards and Recognition:** NeurIPS Scholar Award (2022, 2023); CVPR DEI Award (2023); **Outstanding Reviewer**, CVPR 2025; EarthDial VLM recognized as a *Finalist* in the Scientific Research category at the **UAE AI Award 2025**.
- **Fellowship:** PhD Fellowship, Information Technology University.

Munir, M.A. (last updated: July 2026)

- **Conference Participation:** Presented work at CVPR (2023–2026), NeurIPS 2023, and ICCV 2025.
- **Invited Lectures:** Delivered a hands-on guest lecture in CV805: *Life-long Learning Agents for Vision* at MBZUAI, and an invited guest lecture at the University of Central Florida.
- **Paper Distinctions:** CVPR 2025 [**Highlight**], ICCV 2025 [**Highlight**], ICML 2025 TerraBytes Workshop [**Best Paper**], CVPR 2026 MONTI Workshop [**Best Paper**].

References

References available upon request.