

Career Highlights

- 9100+ citations (H-index:41). 130+ scientific papers in AI/CV venues, including top venues such as, CVPR, ICCV, ECCV, NeurIPS, ICLR, AAAI, IEEE TPAMI, IJCV and IEEE TIP.
- Supervised/co-supervised more than 30 research staff (Post-docs, PhD students, MS students, Visiting students).
- Program Chair of ACM Multimedia Asia'26, Workshop Chair of ACCV'28, Lead Guest Editor of CVIU SI, Workshop Co-organizer at CVPR'24, ICCV'23, NeurIPS'23, ICME'25, Senior Program Committee Member AAAI'26.

Education and Degrees

- 2008 – 2013 **Doctor of Philosophy in Computer Vision**, *Universitat Autònoma de Barcelona, Spain*
- 2007 – 2008 **Master of Science in Computer Vision and Artificial Intelligence**, *Universitat Autònoma de Barcelona, Spain*
- 2005 – 2007 **Master of Science in Intelligent System Design**, *Chalmers University of Technology, Gothenburg, Sweden*

Employment and Professional Experience

- June. 2020 – present **Assistant Professor**, *Mohamed bin Zayed University of Artificial Intelligence, Abu Dhabi, UAE*
- Aug. 2018 – June 2020 **Research Scientist**, *Inception Institute of Artificial Intelligence, Abu Dhabi, UAE*
- Feb. 2014 – July 2018 **Postdoctoral Researcher**, *Aalto University, Helsinki, Finland*
The position was funded under Center of Excellence in Computational Inference Research (COIN).

Research Fundings and Awards

- 2021– 2026 MBZUAI-WIS project titled Deciphering Brain Diseases by AI: Quantifying the Dynamics of Sub-cellular Organelles in Control and Diseased Brain Organoid Models.(Role: *Principal Investigator*, Budget: 800K USD.)
- 2020– present MBZUAI Start-up Fund.
- 2024 MBZUAI Seed Fund for collaboration with Emirates News Agency: WAM-GPT, 2024 (Role: *Principal Investigator*, Budget: 215K AED.)
- 2021– 2026 MBZUAI-WIS project titled Live Imaging and Lineage Tracing of Mouse Embryo Development ex-utero.(Role: *Co-Principal Investigator*, Budget: 800K USD.)
- 2021– 2026 MBZUAI-WIS project titled Biomimetic visual-recognition at low resolution with continuous learning.(Role: *Co-Principal Investigator*, Budget: 800K USD.)
- 2024 Google Research Award project titled A Climate Change and Sustainability Tailored Arabic LLM.(Role: *Co-Principal Investigator*, Budget: 20K USD.)
- 2025-2026 EnergyGPT a collaborative project with MBZUAI IFM, Inception Institute of AI and ADNOC focused on developing long-context multimodal large language model for well-log interpretation and petrophysical reasoning.(Role: *Co-Principal Investigator*, Budget: 400K USD.)
- 2014-2015 PicSOM Experiments in TRECVID from Aalto University, Finland. Member of the PicSOM team which participated in Semantic Indexing (SIN), Multimedia Event Detection (MED) and Multimedia Event Recounting (MER) tasks of TRECVID 2014 and 2015.
- 2012 VOC PASCAL CHALLENGE. Member of the UAB Spain team for the Image Classification task and Detection task. We obtained *Honourable Mention* for Image Classification Challenge task.

Publications

Publications Indices:

Google Scholar Link: https://scholar.google.fi/citations?user=_KlvMVoAAAAJ&hl=en

H-index (Google Scholar): 41

Number of Citations (Google Scholar): > 9138

1. Jinxing Zhou, Yanghao Morty Zhou, Mingfei Han, Tong Wang, Xiaojun Chang, Hisham Cholakkal, **Rao Muhammad Anwer**. "Think before you segment: An object-aware reasoning agent for referring audio-visual segmentation", *Conference on Artificial Intelligence (AAAI)*, Singapore, 2026.
2. Dmitry Demidov, Muhammad Zaigham Zaheer, Zongyan Han, Omkar Thawakar, **Rao Muhammad Anwer**. "Thinking Beyond Labels: Vocabulary-Free Fine-Grained Recognition using Reasoning-Augmented LMMs", *International Conference on Computer Vision and Pattern Recognition (CVPR)*, 2026.
3. Shravan Venkatraman, Ritesh Thawkar, Omkar Thawakar, **Rao Muhammad Anwer**, Hisham Cholakkal, Salman Khan, Fahad Khan. "Paying More Attention to Visual Tokens in Self-Evolving Large Multimodal Models", *European Conference on Computer Vision (ECCV)*, Malmo, Sweden, 2026.
4. Amol Harsh, Zongyan Han, Jean Lahoud, Ye Liu, **Rao Muhammad Anwer**, Hisham Cholakkal, Salman Khan, Fahad Khan. "Ground3D-LMM: Fine-Grained 3D Point Grounding and Spatial Reasoning with LMM", *European Conference on Computer Vision (ECCV)*, Malmo, Sweden, 2026.
5. Shubham Patle, Hania Tariq, Mohammad Usman Khan, Sara Ghaboura, Omkar Thawakar, **Rao Muhammad Anwer**, Salman Khan. "DuwatBench: Bridging Language and Visual Heritage through an Arabic Calligraphy Benchmark for Multimodal Understanding", *Findings of the Association for Computational Linguistics, (ACL)*, 2026.
6. Sara Ghaboura, Ketan More, Wafa Al Ghallabi, Omkar Thawakar, Jorma Laaksonen, Hisham Cholakkal, Salman Khan, **Rao Muhammad Anwer**. "ARB: A Comprehensive Arabic Multimodal Reasoning Benchmark", in *(LREC)*, Palma, Mallorca Spain, 2026.
7. Hazza Mahmood, Yongqiang Yu and **Rao Muhammad Anwer**. "AgriChain: Visually-Grounded Expert-Verified Reasoning for Interpretable Agricultural Vision-Language Models", in *(LREC)* Palma, Mallorca, Spain, 2026,
8. Muhammad Sohail Danish, Muhammad Akhtar Munir, Syed Roshaan 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)*, Rio de Janeiro, Brazil, 2026.
9. Tajamul Ashraf, Amal Saqib, Hanan Ghani, Muhra AlMahri, Yuhao Li, Noor Ahsan, Umair Nawaz, Jean Lahoud, Hisham Cholakkal, Mubarak Shah, Philip Torr, Fahad Shahbaz Khan, **Rao Muhammad Anwer**, Salman Khan. "Agent-X: Evaluating Deep Multimodal Reasoning in Vision-Centric Agentic Tasks", in *International Conference on Learning Representations, (ICLR)*, Rio de Janeiro, Brazil, 2026.
10. Tooba Tehreem Sheikh, Jean Lahoud, **Rao Muhammad Anwer**, Fahad Shahbaz Khan, Salman Khan, Hisham Cholakkal. "MedROV: Towards Real-Time Open-Vocabulary Detection Across Diverse Medical Imaging Modalities", in *Winter Conference on Applications of Computer Vision, (WACV)*, Tucson, USA, 2026.
11. Chao Qin, Fahad Shahbaz Khan, Jiale Cao, Huazhu Fu, Shadab Khan, Rao Muhammad Anwer. "UniNuc- Unified Automatic and Interactive Nucleus Instance Segmentation in Histopathology", *Medical Imaging with Deep Learning (MIDL)*, Taipei, 2026.
12. Sahal Shaji Mullappilly, Mohammed Irfan Kurpath, Jaseel Muhammad Kaithakkodan, Omair Mohamed, Mohamed zidan, Fahad Shahbaz Khan, Salman Khan, **Rao Muhammad Anwer**, Hisham Cholakkal. "MediX-R1: Open Ended Medical Reinforcement Learning", *Findings of the Association for Computational Linguistics, (ACL)*, 2026, under review.
13. Omkar Thawakar, Dmitry Demidov, Vaishnav Potlapalli, Sai Prasanna Teja Reddy Bogireddy, Viswanatha Reddy Gajjala, Alaa Mostafa Lasheen, **Rao Muhammad Anwer**, Fahad Shahbaz Khan. "CoVR-R: Reason-Aware Composed Video Retrieval", *arXiv:2603.20190*, 2026.
14. Jinxing Zhou, Yanghao Zhou, Yaoting Wang, Zongyan Han, Jiaqi Ma, Henghui Ding, **Rao Muhammad Anwer**, Hisham Cholakkal. "Audit After Segmentation: Reference-Free Mask Quality Assessment for Language-Referred Audio-Visual Segmentation", *arXiv:2602.03892*, 2026.
15. Ritesh Thawkar, Shravan Venkatraman, Omkar Thawakar, Abdelrahman Shaker, Fahad Khan, Hisham Cholakkal, Salman Khan, **Rao Muhammad Anwer**. "Ask, Solve, Generate: Self-Evolving Unified Multimodal Understanding and Generation via Self-Consistency Rewards", *arXiv:2606.27376*, 2026.
16. Komal Kumar, **Rao Muhammad Anwer**, Fahad Shahbaz Khan, Salman Khan, Ivan Laptev, Hisham Cholakkal. "DEFT: Decompositional Efficient Fine-Tuning for Text-to-Image Models", *Advances in Neural Information Processing Systems*, 2006.

17. Sara Ghaboura, Ketan Pravin More, Ritesh Thawkar, Wafa Al Ghallabi, Omkar Thawakar, Fahad Shahbaz Khan, Hisham Cholakkal, Salman Khan, **Rao Muhammad Anwer**. "Time travel: A comprehensive benchmark to evaluate llms on historical and cultural artifacts", *Findings of the Association for Computational Linguistics (ACL)*, Vienna, Austria, 2025.
18. Sara Ghaboura, Ahmed Heakl, Omkar Thawakar, Ali Husain Salem Abdulla Alharthi, Ines Riahi, Abduljalil Radman, Jorma Laaksonen, Fahad Shahbaz Khan, Salman Khan, **Rao Muhammad Anwer**. "Camel-bench: A comprehensive arabic llm benchmark", *Findings of the Association for Computational Linguistics (NAACL)*, Albuquerque, New Mexico, 2025.
19. Zongyan Han, Mohamed El Amine Boudjoghra, Jiahua Dong, Jinhong Wang, **Rao Muhammad Anwer**. "All in One: Visual-Description-Guided Unified Point Cloud Segmentation", *International Conference on Computer Vision (ICCV)*, Hawaii, USA, 2025.
20. Wafa Al Ghallabi, Ritesh Thawkar, Sara Ghaboura, Ketan Pravin More, Omkar Thawakar, Hisham Cholakkal, Salman Khan, **Rao Muhammad Anwer**. "Fann or flop: A multigenre, multiera benchmark for arabic poetry understanding in llms", *Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Suzhou, China, 2025.
21. Jinhong Wang, Tajamul Ashraf, Zongyan Han, Jorma Laaksonen, **Rao Muhammad Anwer**. "MIRA: A Novel Framework for Fusing Modalities in Medical RAG", *ACM International Conference on Multimedia (ACM MM)*, Dublin, Ireland, 2025.
22. Muhammad Awais, Ali Husain Salem Abdulla Alharthi, Amandeep Kumar, Hisham Cholakkal, **Rao Muhammad Anwer**. "AgroGPT: Efficient agricultural vision-language model with expert tuning", *Winter Conference on Applications of Computer Vision (WACV)*, Tucson, USA, 2025.
23. Muhammad Awais, Mehaboobathunnisa Sahul Hameed, Bidisha Bhattacharya, Orly Reiner, **Rao Muhammad Anwer**. "BOrg: A Brain Organoid-Based Mitosis Dataset for Automatic Analysis of Brain Diseases", *International Symposium on Biomedical Imaging (ISBI)*, Houston, USA, 2025.
24. Zongyan Han, Jiale Cao, Shuo Chen, Tong Wang, Jorma Laaksonen, **Rao Muhammad Anwer**. "OpenSeg-R: Improving Open-Vocabulary Segmentation via Step-by-Step Visual Reasoning", *arXiv:2505.16974*, 2025.
25. Umair Nawaz, Muhammad Zaigham Zaheer, Fahad Shahbaz Khan, Hisham Cholakkal, Salman Khan, **Rao Muhammad Anwer**. "AI in agriculture: A survey of deep learning techniques for crops, fisheries and livestock", in *arXiv:2507.22101*, 2025.
26. Umair Nawaz, Muhammad Awais, Hanan Gani, Muzammal Naseer, Fahad Shahbaz Khan, Salman Khan, **Rao Muhammad Anwer**. "AgriCLIP: Adapting CLIP for agriculture and livestock via domain-specialized cross-model alignment", *International Conference on Computational Linguistics (COLING)*, Abu Dhabi, UAE, 2025.
27. Ayesha Ishaq, Mohamed El Amine Boudjoghra, Jean Lahoud, Fahad Shahbaz Khan, Salman Khan, Hisham Cholakkal, **Rao Muhammad Anwer**. "Open3dtrack: Towards open-vocabulary 3d multi-object tracking", *International Conference on Robotics and Automation (ICRA)*, Atlanta, USA, 2025.
28. Asif Hanif, Zaigham Zaheer, Salman Khan, Fahad Shahbaz Khan, **Rao Muhammad Anwer**. SPARTA: Spectral Prompt Agnostic Adversarial Attack on Medical Vision-Language Models", *International Workshop on Uncertainty for Safe Utilization of Machine Learning*, 2025.
29. Ayesha Ishaq, Jean Lahoud, Ketan More, Omkar Thawakar, Ritesh Thawkar, Dinura Dissanayake, Noor Ahsan, Yuhao Li, Fahad Shahbaz Khan, Hisham Cholakkal, Ivan Laptev, **Rao Muhammad Anwer**, Salman Khan. "Drivellm-o1: A step-by-step reasoning dataset and large multimodal model for driving scenario understanding", *International Conference on Intelligent Robots and Systems (IROS)*, 2025.
30. Chao Qin, Jiale Cao, Fahad Shahbaz Khan, Salman Khan, Huazhu Fu, Ehud Ahissar, **Rao Muhammad Anwer**. Real-time Breast Lesion Detection in Videos via Spatial-temporal Feature Aggregation", *Medical Imaging with Deep Learning (MIDL)*, Salt Lake, USA, 2025.
31. Ayesha Ishaq, Jean Lahoud, Fahad Shahbaz Khan, Salman Khan, Hisham Cholakkal, **Rao Muhammad Anwer**. "Tracking meets large multimodal models for driving scenario understanding", *British Machine Vision Conference (BMVC)*, Sheffield, UK, 2025.
32. Qi Fan, Kaiqi Liu, Nian Liu, Hisham Cholakkal, **Rao Muhammad Anwer**, Wenbin Li, Yang Gao. "Adapting In-Domain Few-Shot Segmentation to New Domains without Source Domain Retraining", *International Conference on Computer Vision (ICCV)*, Hawaii, USA, 2025.
33. Yevheniia Kryklyvets, Mohammed Irfan Kurpath, Sahal Shaji Mullappilly, Jinxing Zhou, Fahad Shahbaz Khan, **Rao Muhammad Anwer**, Salman Khan, Hisham Cholakkal. "MAViS: A Multimodal Conversational Assistant For Avian Species", *Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Suzhou, China, 2025.
34. Komal Kumar, **Rao Muhammad Anwer**, Fahad Shahbaz Khan, Salman Khan, Ivan Laptev, Hisham Cholakkal.

- "DEFT: Compositional Efficient Fine-Tuning for Text-to-Image Models", *Annual Conference on Neural Information Processing Systems (NeurIPS)*, San Diego, USA, 2025.
35. Dinura Dissanayake, Ahmed Heakl, Omkar Thawakar, Noor Ahsan, Ritesh Thawkar, Ketan More, Jean Lahoud, **Rao Muhammad Anwer**, Hisham Cholakkal, Ivan Laptev, Fahad Shahbaz Khan, Salman Khan. "How Good are Foundation Models in Step-by-Step Embodied Reasoning?", in *arXiv:2509.15293*, 2025.
 36. Yuanwei Liu, Nian Liu, Tao Jiang, Xiwen Yao, **Rao Muhammad Anwer**, Hisham Cholakkal, Junwei Han. "Bridge the Intra-Class Gap: K-shot Multi-scale Intermediate Prototype Mining Transformer for Few-Shot Semantic Segmentation", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (**TPAMI**), 2025.
 37. Sambal Shikhar, Mohammed Irfan Kurpath, Sahal Shaji Mullappilly, Jean Lahoud, Fahad Shahbaz Khan, **Rao Muhammad Anwer**, Salman Khan, Hisham Cholakkal. "LLMVoX: Autoregressive streaming text-to-speech model for any LLM", *Findings of the Association for Computational Linguistics*, (**ACL**), Vienna, Austria, 2025.
 38. Omkar Thawakar, Dinura Dissanayake, Ketan Pravin More, Ritesh Thawkar, Ahmed Heakl, Noor Ahsan, Yuhao Li, Mohammed Zumri, Jean Lahoud, **Rao Muhammad Anwer**, Hisham Cholakkal, Ivan Laptev, Mubarak Shah, Fahad Shahbaz Khan, Salman Khan. "Llamav-o1: Rethinking step-by-step visual reasoning in llms", *Findings of the Association for Computational Linguistics*, (**ACL**), Vienna, Austria, 2025.
 39. Long Li, Huichao Xie, Nian Liu, Dingwen Zhang, **Rao Muhammad Anwer**, Hisham Cholakkal, Junwei Han. "Advanced Discriminative Co-Saliency and Background Mining Transformer for Co-Salient Object Detection", *IEEE Trans. Pattern Anal. Mach. Intell.* (**TPAMI**), 47(9): 7733–7751, 2025.
 40. Tooba Tehreem Sheikh, Komal Kumar, Muhammad Awais, Noa Novershtern, **Rao Muhammad Anwer**, Hisham Cholakkal. *IHA-YOLO: Inter-Head Attention for Real-Time Cell Detection*", *International Symposium on Biomedical Imaging*, (**ISBI**), Houston, USA, 2025.
 41. Komal Kumar, Tajamul Ashraf, Omkar Thawakar, **Rao Muhammad Anwer**, Hisham Cholakkal, Mubarak Shah, Ming-Hsuan Yang, Philip H.S. Torr, Fahad Shahbaz Khan, Salman Khan. "Llm post-training: A deep dive into reasoning large language models", *arXiv:2502.21321*, 2025.
 42. Ahmed Heakl, Sara Ghaboura, Omkar Thawakar, Fahad Shahbaz Khan, Hisham Cholakkal, **Rao Muhammad Anwer**, Salman Khan. "AIN: The arabic inclusive large multimodal model", *arXiv:2502.00094*, 2025.
 43. Vishal Thengane, Jean Lahoud, Hisham Cholakkal, **Rao Muhammad Anwer**, Lu Yin, Xiatian Zhu, Salman Khan. *CLIMB-3D: Continual Learning for Imbalanced 3D Instance Segmentation*", *British Machine Vision Conference*, (**BMVC**), Sheffield, UK, 2025.
 44. Ziyang Luo, Nian Liu, Xuguang Yang, Salman Khan, **Rao Muhammad Anwer**, Hisham Cholakkal, Fahad Shahbaz Khan, Junwei Han. "TAViS: Text-bridged Audio-Visual Segmentation with Foundation Models", *International Conference on Computer Vision (ICCV)*, Hawaii, USA, 2025.
 45. Omkar Thawakar, Dmitry Demidov, Ritesh Thawkar, **Rao Muhammad Anwer**, Mubarak Shah, Fahad Shahbaz Khan, Salman Khan. "Beyond simple edits: Composed video retrieval with dense modifications", *International Conference on Computer Vision (ICCV)*, Hawaii, USA, 2025.
 46. Dongming Wu, Yanping Fu, Saike Huang, Yingfei Liu, Fan Jia, Nian Liu, Feng Dai, Tiancai Wang, **Rao Muhammad Anwer**, Fahad Shahbaz Khan, Jianbing Shen. "RAGNet: Large-scale Reasoning-based Affordance Segmentation Benchmark towards General Grasping", *International Conference on Computer Vision (ICCV)*, Hawaii, USA, 2025.
 47. Ashmal Vayani, Dinura Dissanayake, Hasindri Watawana, Noor Ahsan, Nevasini Sasikumar, Omkar Thawakar, Henok Biadglign Ademtew, Yahya Hmaiti, Amandeep Kumar, Kartik Kuckreja, Mykola Maslych, Wafa Al Ghalabi, Mihail Minkov Mihaylov, Chao Qin, Abdelrahman M. Shaker, Mike Zhang, Mahardika Krisna Ihsani, Amiel Gian Esplana, Monil Gokani, Shachar Mirkin, Harsh Singh, Ashay Srivastava, Endre Hamerlik, Fathinah Asma Izzati, Fadillah Adamsyah Maani, Sebastian Cavada, Jenny Chim, Rohit Gupta, Sanjay Manjunath, Kamila Zhumakhanova, Feno Heriniaina Rabevohitra, Azril Hafizi Amirudin, Muhammad Ridzuan, Daniya Najiha Abdul Kareem, Ketan Pravin More, Kunyang Li, Pramesh Shakya, Muhammad Saad, Amirpouya Ghasemaghahi, Amirbek Djanibekov, Dilshod Azizov, Branislava Jankovic, Naman Bhatia, Alvaro Cabrera, Johan S. Obando-Ceron, Olympiah Otieno, Fabian Farestam, Muztoba Rabbani, Sanoojan Baliah, Santosh Sanjeev, Abduragim Shtanchaev, Maheen Fatima, Thao Nguyen, Amrin Kareem, Toluwani Aremu, Nathan Augusto Zacarias Xavier, Amit Bhatkal, Hawau Olamide Toyin, Aman Chadha, Hisham Cholakkal, **Rao Muhammad Anwer**, Michael Felsberg, Jorma Laaksonen, Tamar Solorio, Monojit Choudhury, Ivan Laptev, Mubarak Shah, Salman Khan, Fahad Shahbaz Khan. *All languages matter: Evaluating llms on culturally diverse 100 languages*", *International Conference on Computer Vision and Pattern Recognition (CVPR)*, Nashville, USA, 2025.
 48. Muhammad Awais, Muzammal Naseer, Salman Khan, **Rao Muhammad Anwer**, Hisham Cholakkal, Mubarak Shah, Ming-Hsuan Yang, Fahad Shahbaz Khan. "Foundational Models Defining a New Era in Vision: A Survey and Outlook", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (**TPAMI**), 2025.

49. Omkar Thawakar, Sanath Narayan, Hisham Cholakkal, **Rao Muhammad Anwer**, Salman Khan, Jorma Laaksonen, Mubarak Shah, Fahad Shahbaz Khan. "Video Instance Segmentation in an Open-World", *International Journal of Computer Vision (IJCV)*, vol 133(1), 398-409, 2025.
50. Muhammad Maaz, Hanoona Rasheed, Abdelrahman Shaker, Salman Khan, Hisham Cholakkal, **Rao Muhammad Anwer**, Tim Baldwin, Michael Felsberg, Fahad Shahbaz Khan. "PALO: A Polyglot Large Multimodal Model for 5B People", *Proceedings of Winter Conference on Applications of Computer Vision (WACV)*, Tucson, USA, 2025.
51. Omkar Thawakar, Ashmal Vayani, Salman Khan, Hisham Cholakkal, **Rao Muhammad Anwer**, Michael Felsberg, Tim Baldwin, Eric P Xing, Fahad Shahbaz Khan. "MobiLLama: Towards Accurate and Lightweight Fully Transparent GPT", *International Conference on Learning Representations SLLM Workshop (ICLR)*, Singapore, 2025.
52. Mohamed El Amine Boudjoghra, Angela Dai, Jean Lahoud, Hisham Cholakkal, **Rao Muhammad Anwer**, Salman Khan, Fahad Shahbaz Khan. "Open-YOLO 3D: Towards Fast and Accurate Open-Vocabulary 3D Instance Segmentation", *International Conference on Learning Representations (ICLR)*, Singapore, 2025.
53. Ge Li, Jiale Cao, Hanging Sun, **Rao Muhammad Anwer**, Jin Xie, Fahad Shahbaz Khan, Yanwei Pang. "Video Instance Segmentation Without Using Mask and Identity Supervision", *IEEE Trans. Multim. (TMM)*, 27: 224-235, 2025.
54. Yuanwei Liu, Nian Liu, Yi Wu, Hisham Cholakkal, **Rao Muhammad Anwer**, Xiwen Yao, Jinwei Han. "Ntrent++: Unleashing the power of non-target knowledge for few-shot semantic segmentation", *IEEE Trans. Circuits Syst. Video Technol. (TCVT)*, 35(5): 4314-4328, 2025.
55. Sahal Shaji Mullappilly, Mohammed Irfan Kurpath, Sara Pieri, Saeed Yahya Alseiari, Shanavas Cholakkal, Khaled Aldahmani, Fahad Shahbaz Khan, **Rao Muhammad Anwer**, Salman Khan, Timothy Baldwin, Hisham Cholakkal, "BiMediX2: Bio-Medical EXpert LMM for Diverse Medical Modalities", *Findings of the Association for Computational Linguistics (EMNLP)*, Suzhou, China, 2025.
56. Yuanwei Liu, Junwei Han, Xiwen Yao, Salman Khan, Hisham Cholakkal, **Rao Muhammad Anwer**, Nian Liu, Fahad Shahbaz Khan. "Bidirectional Reciprocal Information Communication for Few-Shot Semantic Segmentation." *Forty-first International Conference on Machine Learning (ICML)*, Vienna, Austria, 2024.
57. Mubashir Noman, Muzammal Naseer, Hisham Cholakkal, **Rao Muhammad Anwer**, Salman Khan, Fahad Shahbaz Khan. "Rethinking transformers pre-training for multi-spectral satellite imagery", *International Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, USA, 2024.
58. Omkar Thawakar, Muzammal Naseer, **Rao Muhammad Anwer**, Salman Khan, Michael Felsberg, Mubarak Shah, Fahad Shahbaz Khan. "Composed Video Retrieval via Enriched Context and Discriminative Embeddings", *International Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, USA, 2024.
59. Hanoona Rasheed, Muhammad Maaz, Sahal Shaji, Abdelrahman Shaker, Salman Khan, Hisham Cholakkal, **Rao Muhammad Anwer**, Erix Xing, Ming-Hsuan Yang, Fahad S Khan. "GLaMM: Pixel Grounding Large Multimodal Model", *International Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, USA, 2024.
60. Long Li, Junwei Han, Nian Liu, Salman Khan, Hisham Cholakkal, **Rao Muhammad Anwer**, Fahad Shahbaz Khan. "Robust Perception and Precise Segmentation for Scribble-Supervised RGB-D Saliency Detection", *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 46(1): 479-496, 2024.
61. Mustansar Fiaz, Mubashir Noman, Hisham Cholakkal, **Rao Muhammad Anwer**, Jacob Hanna, Fahad Shahbaz Khan. "Guided-attention and gated-aggregation network for medical image segmentation", *Pattern Recognit. (PR)* 156: 110812, 2024.
62. Bin Xie, Jiale Cao, **Rao Muhammad Anwer**, Jin Xie, Jing Nie, Aiping Yang, Yanwei Pang. "Multi-query and multi-level enhanced network for semantic segmentation". *Pattern Recognit. (PR)* 156:110777, 2024.
63. Lei Huang, Yunhao Ni, Xi Weng, **Rao Muhammad Anwer**, Salman Khan, Ming-Hsuan Yang, Fahad Shahbaz Khan. "Understanding Whitening Loss in Self-supervised Learning", *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 2024.
64. Amandeep Kumar, Muhammad Awais, Sanath Narayan, Hisham Cholakkal, Salman Khan, **Rao Muhammad Anwer**. "Efficient 3D-Aware Facial Image Editing via Attribute-Specific Prompt Learning", *European Conference on Computer Vision (ECCV)*, Milano, Italy, 2024.
65. Long Li, Junwei Han, Dingwen Zhang, Zhongyu Li, Salman Khan, **Rao Muhammad Anwer**, Hisham Cholakkal, Nian Liu, Fahad Shahbaz Khan. "CONDA: Condensed Deep Association Learning for Co-Salient Object Detection", *European Conference on Computer Vision (ECCV)*, Milano, Italy, 2024.
66. Mohamed El Amine Boudjoghra, Jean Lahoud, Salman Khan, Hisham Cholakkal, **Rao Muhammad Anwer**, Fahad Shahbaz Khan. "OpenDistill3D: Open-World 3D Instance Segmentation with Unified Self-Distillation for Continual Learning and Unknown Class Discovery", *European Conference on Computer Vision (ECCV)*, Milano,

Italy, 2024.

67. Mubashir Noman, Noor Ahsan, Muzammal Naseer, Hisham Cholakkal, **Rao Muhammad Anwer**, Salman Khan, Fahad Shahbaz Khan. "CDChat: A Large Multimodal Model for Remote Sensing Change Description", in *arxiv.org/abs/2409.16261*, 2024.
68. Binglu Wang, Haowen Zheng, Lei Zhang, Nian Liu, **Rao Muhammad Anwer**, Hisham Cholakkal, Yongqiang Zhao, Zhijun Li. "BEVRefiner: Improving 3D Object Detection in Bird's-Eye-View via Dual Refinement", *IEEE Transactions on Intelligent Transportation Systems*, **TITS** 25(10):15094–15105, 2024.
69. Mubashir Noman, Mustansar Fiaz, Hisham Cholakkal, Sanath Narayan, **Rao Muhammad Anwer**, Salman Khan, Fahad Shahbaz Khan. "Remote sensing change detection with transformers trained from scratch", *IEEE Transactions on Geoscience and Remote Sensing* (2024).
70. Asif Hanif, Fahad Shamshad, Muhammad Awais, Muzammal Naseer, Fahad Shahbaz Khan, Karthik NanadaKumar, Salman Khan, **Rao Muhammad Anwer**. "BAPLe: Backdoor Attacks on Medical Foundational Models using Prompt Learning", *International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI)*, Marrakesh, Morocco, 2024.
71. Chao Qin, Jiale Cao, Huazhu Fu, **Rao Muhammad Anwer**, Fahad Shahbaz Khan. "DB-SAM: Delving into High Quality Universal Medical Image Segmentation", *International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI)*, Marrakesh, Morocco, 2024. **Best Paper Finalist**.
72. Sahal Shaji Mullappilly, Abhishek Singh Gehlot, **Rao Muhammad Anwer**, Fahad Shahbaz Khan, Hisham Cholakkal. "Semi-supervised Open-World Object Detection", *Conference on Artificial Intelligence (AAAI)*, Vancouver, Canada, 2024.
73. Sara Pieri, Sahal Shaji Mullappilly, Fahad Shahbaz Khan, **Rao Muhammad Anwer**, Salman Khan, Timothy Baldwin, Hisham Cholakkal. "BiMediX: Bilingual Medical Mixture of Experts LLM." *Findings of the Association for Computational Linguistics (EMNLP)*, Florida, USA, 2024.
74. Amandeep Kumar, Muzammal Naseer, Sanath Narayan, **Rao Muhammad Anwer**, Salman Khan, Hisham Cholakkal. "Multi-modal Generation via Cross-Modal In-Context Learning." in *arXiv:2405.18304* (2024).
75. Xi Weng, Yunhao Ni, Tengwei Song, Jie Luo, **Rao Muhammad Anwer**, Salman Khan, Fahad Shahbaz Khan, Lei Huang. "Modulate Your Spectrum in Self-Supervised Learning", *International Conference on Learning Representations (ICLR)*, Vienna, Austria, 2024.
76. Omkar Thawkar, Abdelrahman Shaker, Sahal Shaji Mullappilly, Hisham Cholakkal, **Rao Muhammad Anwer**, Salman Khan, Jorma Laaksonen, Fahad Shahbaz Khan. "XrayGPT: Chest Radiographs Summarization using Medical Vision-Language Models", *Workshop on Biomedical Natural Language Processing, (BioNLP)*, Bangkok, Thailand, 2024.
77. Long Li, Junwei Han, Ni Zhang, Nian Liu, Salman Khan, Hisham Cholakkal, **Rao Muhammad Anwer**, Fahad Shahbaz Khan. "Discriminative Co-Saliency and Background Mining Transformer for Co-Salient Object Detection", *International Conference on Computer Vision and Pattern Recognition (CVPR)*, Vancouver, Canada, 2023.
78. Ankan Kumar Bhunia, Salman Khan, Hisham Cholakkal, **Rao Muhammad Anwer**, Jorma Laaksonen, Mubarak Shah, Fahad Shahbaz Khan. "Person Image Synthesis via Denoising Diffusion Model", *International Conference on Computer Vision and Pattern Recognition (CVPR)*, Vancouver, Canada, 2023.
79. Amandeep Kumar, Ankan Kumar Bhunia, Sanath Narayan, Hisham Cholakkal, **Rao Muhammad Anwer**, Salman Khan, Ming-Hsuan Yang, Fahad Shahbaz Khan. "Generative Multiplane Neural Radiance for 3D-Aware Image Generation", *International Conference on Computer Vision (ICCV)*, Paris, France, 2023.
80. Nian Liu, Kepan Nan, Wangbo Zhao, Yuanwei Liu, Xiwen Yao, Salman Khan, Hisham Cholakkal, **Rao Muhammad Anwer**, Junwei Han, Fahad Shahbaz Khan. "Multi-grained Temporal Prototype Learning for Few-shot Video Object Segmentation", *International Conference on Computer Vision (ICCV)*, Paris, France, 2023.
81. Mohamed El Amine Boudjoghra, Salwa K Al Khatib, Jean Lahoud, Hisham Cholakkal, **Rao Muhammad Anwer**, Salman Khan, Fahad Khan. "3D Indoor Instance Segmentation in an Open-World", *Annual Conference on Neural Information Processing Systems (NeurIPS)*, New Orleans, USA, 2023.
82. Omkar Thawakar, **Rao Muhammad Anwer**, Jorma Laaksonen, Orly Reiner, Mubarak Shah, Fahad Shahbaz Khan. "3D Mitochondria Instance Segmentation with Spatio-Temporal Transformers", *International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI)*, Vancouver, Canada, 2023.
83. Chao Qin, Jiale Cao, Huazhu Fu, **Rao Muhammad Anwer**, Fahad Shahbaz Khan. "A Spatial-Temporal Deformable Attention Based Framework for Breast Lesion Detection in Videos", *International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI)*, Vancouver, Canada, 2023.
84. Amandeep Kumar, Ankan Kumar Bhunia, Sanath Narayan, Hisham Cholakkal, **Rao Muhammad Anwer**, Jorma Laaksonen, Fahad Shahbaz Khan. "Cross-Modulated Few-Shot Image Generation for Colorectal Tissue Classifi-

- cation", *International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI)*, Vancouver, Canada, 2023.
85. Mustansar Fiaz, Hisham Cholakkal, **Rao Muhammad Anwer**, Fahad Shahbaz Khan. "SAT: Scale-Augmented Transformer for Person Search" *Proceedings of Winter Conference on Applications of Computer Vision (WACV)*, Waikoloa, USA, 2023.
 86. Jiale Cao, Yanwei Pang, **Rao Muhammad Anwer**, Hisham Cholakkal, Fahad Shahbaz Khan, Ling Shao. "Sip-Maskv2: Enhanced Fast Image and Video Instance Segmentation", *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 45(3): 3798–3812, 2023.
 87. Abdulaziz Amer Aleissae, Amandeep Kumar, **Rao Muhammad Anwer**, Salman Khan, Hisham Cholakkal, Gui-Song Xia. "Transformers in Remote Sensing: A Survey", *Remote. Sensing*, 15(7): 1860, 2023.
 88. Reem Abdalla Sharif, Mustansar Fiaz, **Rao Muhammad Anwer**. "PSM-PS: Part-Based Signal Modulation for Person Search", *International Conference on Computer Analysis of Images and Patterns (CAIP)*, Limassol, Cyprus, 2023.
 89. Mustansar Fiaz, **Rao Muhammad Anwer**, Hisham Cholakkal. "SA2-Net: Scale-aware Attention Network for Microscopic Image Segmentation", *British Machine Vision Conference, (BMVC)*, Aberdeen, UK, 2023.
 90. Zhiqiang Dong, Jiale Cao, **Rao Muhammad Anwer**, Jin Xie, Fahad Khan, Yanwei Pang. "LEAPS: End-to-End One-Step Person Search With Learnable Proposals", in *arXiv:2303.11859*, 2023.
 91. Hefeng Wang, Jiale Cao, **Rao Muhammad Anwer**, Jin Xie, Fahad Shahbaz Khan, Yanwei Pang. "DFormer: Diffusion-guided Transformer for Universal Image Segmentation", in *arXiv:2306.03437*. 2023.
 92. Sultan Abu Ghazal, Jean Lahoud, **Rao Muhammad Anwer**. "Surface-Biased Multi-Level Context 3D Object Detection", *Proceedings of International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, (VISIGRAPP)*, Lisbon, Portugal, 2023.
 93. Sahal Shaji Mullappilly, Abdelrahman Shaker, Omkar Thawakar, Hisham Cholakkal, Rao Muhammad Anwer, Salman Khan, Fahad Shahbaz Khan. "Arabic Mini-ClimateGPT : A Climate Change and Sustainability Tailored Arabic LLM", *Findings of the Association for Computational Linguistics (EMNLP)*, Singapore, 2023.
 94. Xi Weng, Lei Huang, Lei Zhao, Rao Muhammad Anwer, Salman Khan, Fahad Shahbaz Khan. "An Investigation into Whitening Loss for Self-supervised Learning", *Annual Conference on Neural Information Processing Systems (NeurIPS)*, New Orleans, USA, 2022.
 95. Jiale Cao, Yenwai Pang, **Rao Muhammad Anwer**, Hisham Cholakkal, Jin Xie, Mubarak Shah, Fahad Shahbaz Khan. "PSTR: End-to-End One-Step Person Search With Transformers", *International Conference on Computer Vision and Pattern Recognition (CVPR)*, New Orleans, LA, USA, 2022.
 96. K. J. Joseph, Salman Khan, Fahad Shahbaz Khan, **Rao Muhammad Anwer**, Vineeth N Balasubramanian. "Energy-based Latent Aligner for Incremental Learning". *International Conference on Computer Vision and Pattern Recognition (CVPR)*, New Orleans, LA, USA, 2022.
 97. Anirudh Thatipelli, Sanath Narayan, Salman Khan, **Rao Muhammad Anwer**, Fahad Shahbaz Khan, Bernard Ghanem. "Spatio-temporal Relation Modeling for Few-shot Action Recognition". *International Conference on Computer Vision and Pattern Recognition (CVPR)*, New Orleans, LA, USA, 2022.
 98. Dhanalaxmi Gaddam, Jean Lahoud, Fahad Shahbaz Khan, **Rao Muhammad Anwer**, Hisham Cholakkal. "CMR3D: Contextualized Multi-Stage Refinement for 3D Object Detection", *Proceedings of ACM International Conference on Multimedia in Asia, (MMAsia)*, Tokyo, Japan, 2022.
 99. Fatima Albreiki, Sultan Abu Ghazal, Jean Lahoud, **Rao Muhammad Anwer**, Hisham Cholakkal, Fahad Shahbaz Khan. "On the Robustness of 3D Object Detectors", *Proceedings of ACM International Conference on Multimedia in Asia, (MMAsia)*, Tokyo, Japan, 2022.
 100. Muhammad Maaz, Abdelrahman Shaker, Hisham Cholakkal, Salman Khan, Syed Waqas Zamir, **Rao Muhammad Anwer**, Fahad Shahbaz Khan. "EdgeNeXt: Efficiently Amalgamated CNN-Transformer Architecture for Mobile Vision Applications", *European Conference on Computer Vision (ECCV)*, Tel Aviv, Israel, 2022.
 101. Muhammad Maaz, Hanoona Rasheed, Salman Khan, Fahad Shahbaz Khan, **Rao Muhammad Anwer**, Ming-Hsuan Yang. "Class-Agnostic Object Detection with Multi-modal Transformer", *European Conference on Computer Vision (ECCV)*, Tel Aviv, Israel, 2022.
 102. Ankan Kumar Bhunia, Salman Khan, Hisham Cholakkal, **Rao Muhammad Anwer**, Fahad Shahbaz Khan, Jorma Laaksonen, Michael Felsberg. "DoodleFormer: Creative Sketch Drawing with Transformers". *European Conference on Computer Vision (ECCV)*, Tel Aviv, Israel, 2022.
 103. Omkar Thawakar, Sanath Narayan, Jiale Cao, Hisham Cholakkal, **Rao Muhammad Anwer**, Muhammad Haris Khan, Salman Khan, Michael Felsberg, Fahad Shahbaz Khan. "Video Instance Segmentation via Multi-Scale Spatio-Temporal Split Attention Transformer", *European Conference on Computer Vision (ECCV)*, Tel Aviv, Israel, 2022.

104. Jin Xie, **Rao Muhammad Anwer**, Hisham Cholakkal, Jing Nie, Jiale Cao, Jorma Laaksonen, Fahad Shahbaz Khan. "Learning a Dynamic Cross-Modal Network for Multispectral Pedestrian Detection", *International Conference on Multimedia (ACMM)*, Lisboa, Portugal, 2022.
105. Xiaoheng Jiang, Xinyi Wu, Hisham Cholakkal, **Rao Muhammad Anwer**, Jiale Cao Mingliang Xu, Bing Zhou, Yanwei Pang, Fahad Shahbaz Khan. "Multi-scale Feature Aggregation for Crowd Counting", *arXiv preprint arXiv:2208.05256*, 2022.
106. Jean Lahoud, Jiale Cao, Fahad Shahbaz Khan, Hisham Cholakkal, **Rao Muhammad Anwer**, Salman Khan, Ming-Hsuan Yang. "3D Vision with Transformers: A Survey", in *arXiv:2208.04309*, 2022.
107. Mustansar Fiaz, Hisham Cholakkal, Sanath Narayan, **Rao Muhammad Anwer**, Fahad Shahbaz Khan. "PS-ARM: An End-to-End Attention-aware Relation Mixer Network for Person Search", *Proceedings of the Asian Conference on Computer Vision (ACCV)*, Macao, China, 2022.
108. Shivam Chandhok, Sanath Narayan, Hisham Cholakkal, **Rao Muhammad Anwer**, Vineeth N Balasubramanian, Fahad Shahbaz Khan, Ling Shao. "Structured Latent Embeddings for Recognizing Unseen Classes in Unseen Domains", *British Machine Vision Conference, (BMVC)*, Virtual, 2021.
109. Ankan Kumar Bhunia, Salman Khan, Hisham Cholakkal, **Rao Muhammad Anwer**, Fahad Shahbaz Khan, Mubarak Shah. "Handwriting Transformer", *International Conference on Computer Vision (ICCV)*, Virtual, 2021.
110. Jin Xie, Yanwei Pang, Muhammad Haris Khan, **Rao Muhammad Anwer**, Fahad Shahbaz Khan, Ling Shao. "Mask-Guided Attention Network and Occlusion-Sensitive Hard Example Mining for Occluded Pedestrian Detection", *IEEE Transactions on Image Processing (TIP)* 24: 3872-3884, 2021.
111. Jin Xie, Yanwei Pang, Hisham Cholakkal, **Rao Muhammad Anwer**, Fahad Shahbaz Khan, Ling Shao. "PSC-Net: learning part spatial co-occurrence for occluded pedestrian detection", *Science China Information Sciences* 64(2), 2021.
112. **Rao Muhammad Anwer**, Fahad Shahbaz Khan, Jorma Laaksonen. "Compact Deep Color Features for Remote Sensing Scene Classification", *Neural Processing Letters (NPL)* 53(2): 1523-1544, 2021.
113. Jiale Cao, **Rao Muhammad Anwer**, Hisham Cholakkal, Fahad Shahbaz Khan, Yanwei Pang, Ling Shao, Mubarak Shah. "SipMask:Spatial Information Preservation for Fast Image and Video Instance Segmentation." *European Conference on Computer Vision(ECCV)*, Glasgow, Scotland, 2020.
114. Jin Xie, Hisham Cholakkal, **Rao Muhammad Anwer**, Fahad Shahbaz Khan, Yanwei Pang, Ling Shao, Mubarak Shah. "Count-and Similarity-aware R-CNN for Pedestrian Detection." *European Conference on Computer Vision (ECCV)*, Glasgow, Scotland, 2020.
115. Jiale Cao, Hisham Cholakkal, **Rao Muhammad Anwer**, Fahad Shahbaz Khan, Yanwei Pang, Ling Shao. "D2Det: Towards High Quality Object Detection and Instance Segmentation." *International Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, USA, 2020.
116. Jing Nie, **Rao Muhammad Anwer**, Hisham Cholakkal, Fahad Shahbaz Khan, Yanwei Pang, Ling Shao. "Enriched Feature Guided Refinement Network for Object Detection." *International Conference on Computer Vision (ICCV)*, Seoul, Korea, 2019.
117. Tiancai Wang, **Rao Muhammad Anwer**, Hisham Cholakkal, Fahad Shahbaz Khan, Yanwei Pang, Ling Shao. "Learning Rich Features at High-Speed for Single Shot Object Detection." *International Conference on Computer Vision (ICCV)*, Seoul, Korea, 2019.
118. Tiancai Wang, **Rao Muhammad Anwer**, Muhammad Haris Khan, Fahad Shahbaz Khan, Yanwei Pang, Ling Shao, Jorma Laaksonen. "Deep Contextual Attention for Human-Object Interaction Detection." *International Conference on Computer Vision (ICCV)*, Seoul, Korea, 2019.
119. Yanwei Pang, Jin Xie, Muhammad Haris Khan, **Rao Muhammad Anwer**, Fahad Shahbaz Khan, Ling Shao. "Mask-Guided Attention Network for Occluded Pedestrian Detection." *International Conference on Computer Vision (ICCV)*, Seoul, Korea, 2019.
120. Yanwei Pang, Tiancai Wang, **Rao Muhammad Anwer**, Fahad Shahbaz Khan, Ling Shao. "Efficient Featurized Image Pyramid Network for Single Shot Detector." *International Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach USA, 2019.
121. **Rao Muhammad Anwer**, Fahad Shahbaz Khan, Jorma Laaksonen, Nazar Zaki. "Multi-Stream Convolutional Networks for Indoor Scene Recognition." *International Conference on Computer Analysis of Images and Patterns (CAIP)*, Salerno, Italy, 2019. **(Oral)**
122. **Rao Muhammad Anwer**, Fahad Shahbaz Khan, Joost van de Weijer, Matthieu Molinier, Jorma Laaksonen. "Binary Patterns Encoded Convolutional Neural Networks for Texture Recognition and Remote Sensing Scene Classification." *ISPRS Journal of Photogrammetry and Remote Sensing (PHOTO)*, vol 138, 74-85, 2018.
123. Fahad Shahbaz Khan, Joost van de Weijer, **Rao Muhammad Anwer**, Andrew Bagdanov, Michael Felsberg,

- Jorma Laaksonen. "Scale coding bag of deep features for human attribute and action recognition." *Machine Vision and Applications (MVA)*, vol 29(1), 55-71, 2018.
124. Hamed Tavakoli, Ali Borji, **Rao Muhammad Anwer**, Esa Rahtu, Juho Kannala. "Bottom-Up Attention Guidance for Recurrent Image Recognition." *International Conference on Images Processing (ICIP)*, Athens, Greece, 2018.
 125. **Rao Muhammad Anwer**, Fahad Khan, Jorma Laaksonen. "Two-Stream Part-based Deep Representation for Human Attribute Recognition." *IEEE International Conference on Biometrics (ICB)*, Gold Coast, Australia, 2018. (Oral)
 126. **Rao Muhammad Anwer**, Fahad Shahbaz Khan, Joost Van de Weijer, Jorma Laaksonen. "TEX-Nets: Binary Patterns Encoded Convolutional Neural Networks for Texture Recognition." *ACM International Conference on Multimedia Retrieval (ICMR)*, Bucharest, Romania, 2017. (Oral)
 127. **Rao Muhammad Anwer**, Fahad Shahbaz Khan, Joost Van de Weijer, Jorma Laaksonen. "Top-Down Deep Appearance Attention for Action Recognition." *Scandinavian Conference on Image Analysis (SCIA)*, Tromsø, Norway, 2017. (Oral)
 128. **Rao Muhammad Anwer**, Fahad Shahbaz Khan, Joost van de Weijer, Jorma Laaksonen. "Combining Holistic and Part-based Deep Representations for Computational Painting Categorization." *ACM International Conference on Multimedia Retrieval (ICMR)*, New York, USA, 2016.
 129. Fahad Shahbaz Khan, **Rao Muhammad Anwer**, Joost Van de Weijer, Michael Felsberg, Jorma Laaksonen. "Deep Semantic Pyramids for Human Attributes and Action Recognition." *Scandinavian Conference on Image Analysis (SCIA)*, Copenhagen, Denmark, 2015.
 130. Fahad Shahbaz Khan, **Rao Muhammad Anwer**, Joost Van de Weijer, Michael Felsberg, Jorma Laaksonen. "Compact Color-Texture Description for Texture Classification." *Pattern Recognition Letters (PRL)*, vol 51(1), 16-22, 2015.
 131. Fahad Shahbaz Khan, Jiaolong Xu, Joost van de Weijer, Andrew Bagdanov, **Rao Muhammad Anwer**, Antonio Lopez. "Recognizing Actions through Action-Specific Person Detection." *IEEE Transactions on Image Processing (TIP)* 24(11): 4422-4432, 2015.
 132. Fahad Shahbaz Khan, Joost Van de Weijer, **Rao Muhammad Anwer**, Michael Felsberg, Carlo Gatta. "Semantic Pyramids for Gender and Action Recognition." *IEEE Transactions on Image Processing (TIP)*, vol 23(8), 3633-3645, 2014.
 133. Fahad Shahbaz Khan, **Rao Muhammad Anwer**, Joost Van de Weijer, Andrew Bagdanov, Antonio Lopez, Michael Felsberg. "Coloring Action Recognition in Still Images." *International Journal of Computer Vision (IJCV)*, vol 105(3), 205-221, 2013.
 134. Fahad Shahbaz Khan, **Rao Muhammad Anwer**, Joost van de Weijer, Andrew Bagdanov, Maria Vanrell, Antonio Lopez. "Color Attributes for Object Detection." *International Conference on Computer Vision and Pattern Recognition (CVPR)*, Providence, USA, 2012.
 135. **Rao Muhammad Anwer**, David Vázquez, Antonio M López. "Color Contribution to Part-Based Person Detection in Different Types of Scenarios." *International Conference on Computer Analysis of Images and Patterns (CAIP)*, Seville, Spain, 2011.
 136. **Rao Muhammad Anwer**, David Vázquez, Antonio M López. "Opponent Color for Human Detection." *Iberian Conference on Pattern Recognition and Image Analysis (ibPRIA)*, Las Palmas de Gran Canaria, Spain, 2011.

Patents

- System and Method for Video Instance Segmentation Via Multi-Scale Spatio-Temporal Split Attention Transformer, **US Patent Granted Num.: US 12,499,573, 2025.**
- System and method for attention-aware relation mixer for person search, **US Patent Granted, Num.:12260674, 2025.**
- System and Method for Efficiently Amalgamated CNN Transformer Architecture for Mobile Vision Applications. **US Patent Granted, Num.:12373672, 2025.**
- System and Method of Open-World Semi-Supervised Satellite Object Detection. **US Patent Granted, Num.:12380677, 2025.**
- System and method for handwriting generation, **US Patent Granted, Num.:11756244, 2023.**
- Dense and Discriminative Neural Network Architectures for Improved Object Detection and Instance Segmentation, **US Patent Granted, Num.:11244188, 2022.**
- Bilingual medical mixture of experts large language model. **US Patent App. Num. 18/905,019, 2025.**
- LLMVoX: Autoregressive Streaming Text-to-Speech Model for Any LLM. **US Patent App. Num. 63/777,31 , 2025.**
- Method and System for Performing Three Dimensional (3D)-Aware Image Editing. **US Patent App. Num.**

19065624, 2025.

- System and method of cross-modulated dense local fusion for few-shot image generation, **US Patent App. Num. 17983952, 2024.**

RESEARCH SUPERVISION

Post-doc Supervision

- 2020 – present Jean Lahoud. MBZ University of AI (MBZUAI, UAE).
- 2024 – present Zongyan Han. MBZ University of AI (MBZUAI, UAE).
- 2025 – present Jinxing Zhou . MBZ University of AI (MBZUAI, UAE).
- 2023 – 2024 Muhammad Awais. MBZ University of AI (MBZUAI, UAE).
- 2022 – 2023 Mustansar Fiaz. MBZ University of AI (MBZUAI, UAE).

Ph.D. Supervision/Co-supervision

- 2024 – 2026 Sara Ghaboura. MBZ University of AI, UAE.
- 2021 – 2024 Mubashir Noman. MBZ University of AI, UAE.
- 2023 – 2026 Omkar Thawker. MBZ University of AI, UAE.
- 2022 – 2025 Chao Qin. MBZ University of AI, UAE.
- 2023 – 2026 Wafa Hamad Mohamed Ali Alghallabi. MBZ University of AI, UAE.
- 2023 – 2026 Mohammad Ahmad Eid Mohamed Almansoori. MBZ University of AI, UAE.
- 2023 – 2026 Muhammad Ibraheem Siddiqui. MBZ University of AI, UAE.

Master (MSc) Thesis Supervision/co-supervision

- 2020 – 2022 Sultan Abu Ghazal. MBZ University of AI, UAE.
- 2020 – 2022 Fatima Abdulla Omar Alberiki. MBZ University of AI, UAE.
- 2021 – 2023 Reem Abdalla Sharif Mohamed Alameeri. MBZ University of AI, UAE.
- 2021 – 2023 Abdulaziz Amer Mohammed Aleissae, MBZ University of AI, UAE.
- 2020 – 2023 Imane Hilal. MBZ University of AI, UAE.
- 2022 – 2024 Mehoobathunnisa Shul Hameed. MBZ University of AI, UAE.
- 2022 – 2024 Ali Husain Salem Abdulla Alharthi. MBZ University of AI, UAE.
- 2023 – 2025 Umair Nawaz. MBZ University of AI, UAE.
- 2023 – 2025 Tooba Tehreem Sheikh. MBZ University of AI, UAE.
- 2024 – 2026 Hazza Mahmood . MBZ University of AI, UAE.
- 2025 – 2027 Ketan More. MBZ University of AI, UAE.
- 2025 – 2027 Ritesh Thawkar. MBZ University of AI, UAE.
- 2025 – 2027 Abdulla Ibrahim Marzooq Ibrahim Alshehhi. MBZ University of AI, UAE.

Research Intern/Visiting students Supervision/co-supervision

- 2024 – 2025 Tajamul Ashraf . Research Assistant at MBZUAI.
- 2025 – 2026 Xinyu Yan . From Tianjin University, China.
- 2024 – 2025 Dongming Wu. From Beijing Institute of Technology, China.
- 2018 – 2019 Tiancai Wang. From Tianjin University, China.
- 2018 – 2019 Jin Xie. From Tianjin University, China.
- 2018 – 2019 Jing Nie. From Tianjin University, China.
- 2022 – 2024 Amandeep Kumar. Research Assistant , MBZUAI, UAE.
- 2020 – 2021 Shivam Chandhok. Research Assistant, MBZUAI, UAE.

Teaching

During my time at MBZUAI, I have contributed in developing the course material of *multiple courses* from scratch. In addition, I have been course examiner/coordinator of *two* fundamental and large courses offered to *all* the students (CV, NLP, ML) of the University.

- 2021 (Spring-Fall) **Research Communication and Dissemination, MBZUAI, Role:** Course Examiner and Instructor
Responsibilities: Designed the course and served as both examiner and instructor. This new course was prepared from scratch at the inception of the University. The course was taken by the entire MBZUAI batch (Computer Vision, Machine Learning, and Natural language Processing departments) for both PhD and MSc students during Spring 2021 and Fall 2021.
Course Evaluation: Spring 2021 (Course Rating: 88.5%, Instructor Evaluation: 8.73/10.0).
- 2021–present **Visual Object Recognition and Detection, MBZUAI, Role:** Course Instructor
Responsibilities: Lecturer and design of the course. This new course was prepared from scratch in 2021 and has been offered every year afterwards.
Course Evaluation: Spring 2024 (Course Rating: 8.6/10.0, Instructor Evaluation: 4.5/5.0).
- 2022 **Deep Learning, MBZUAI, Role:** Course Instructor
Responsibilities: This elective course, initially co-developed from scratch in 2022, was offered to students from all three departments (Computer Vision, Machine Learning, and Natural language Processing).
Course Evaluation: Spring 2022 (Course Rating: 85.5%, Instructor Evaluation: 8.18/10.0)
- 2022– 2024 **Foundations of Artificial Intelligence, MBZUAI, Role:** Course Examiner and Instructor
Responsibilities: This course was co-developed from scratch in 2022 within a two-month summer period and was offered to the entire MBZUAI Master's cohort across the Computer Vision, Machine Learning, and Natural Language Processing programs in Fall 2022 and Fall 2023.
Course Evaluation: Fall 2023 (Course Rating: 8.2/10.0, Instructor Evaluation: 4.2/5.0)
- 2025 **Selected Topics in Computer Vision, MBZUAI, Role:** Course Examiner and Instructor
Responsibilities: This elective course was designed and developed from scratch. The course was offered for the first time in Fall 2025 to Master students from the Computer Vision, Machine Learning, Natural Language Processing, and Robotics departments. It covers both classical foundations and cutting-edge topics in Computer Vision, delivered through a combination of in-class lectures, laboratory sessions, and multiple paper discussion and presentation activities. I served both as a course examiner and instructor for this course. The course received highly positive student feedback.
Course Evaluation: Fall 2025 (Course Rating: 5.0/5.0, Instructor Evaluation: 5.0/5.0)
- 2026 **Advanced Topics in Large Multimodal Models, MBZUAI, Role:** Course Examiner and Instructor
Responsibilities: This PhD elective course was developed from scratch and is scheduled for its first offering in Spring 2026 to students from the Computer Vision, Machine Learning, Natural Language Processing, and Robotics programs. The course covers fundamentals of large multimodal models (LMMs), LMMs for discriminative vision tasks (e.g., region and pixel grounding in images and videos), challenges in LMMs such as hallucination and bias (including across languages and cultures), and emerging areas such as LMMs for open-vocabulary perception and visual reasoning. The content is covered through in-class lectures, laboratory sessions, and multiple instructor-led research paper discussion sessions.

Languages

Urdu (native), English (fluent)