Omkar Thawakar

+971-585-25-1761 | thawakar.omkar@gmail.com | omkarthawakar.github.io

in LinkedIn | ○ Github | Ⓖ Scholar

Abu Dhabi, UAE

RESEARCH INTEREST

Passionate researcher in video retrieval, multimodal learning, and large language models (LLMs), with hands-on experience in video instance segmentation, dense retrieval, and large-scale multimodal datasets. Interested in contributing innovative ideas, technical knowledge, and dedication to organizations that foster collaborative research.

EXPERIENCE

- **Mohamed bin Zayed University of Artificial Intelligence (MBZUAI) [⊕]**   *May 2021 - Jul 2023*
  Researcher   Abu Dhabi, UAE
  - Worked on Composed Video Retrieval, where we proposed the idea of using dense video description to retrieve the target video with desired compositional changes.
  - Worked on video instance segmentation with transformers. (SOTA performance, ECCV 2022)
  - Implemented 3D Mitochondria Instance Segmentation with Spatio Temporal Transformers with more than 3% overall improvement in segmentation accuracy over recent best. (SOTA performance, MICCAI-2023)
  - Worked on open-world video instance segmentation where we presented the finding on using shallow features for novel object detection. (IJCV-2023)

- **Chefling India Pvt Ltd [⊕]**   *Feb 2020 - May 2021*
  Machine Learning Engineer   Bangalore, India
  - Worked on creating and deploying own recipe scraping service with Machine Learning model supported by python scrapy tools, increasing the accuracy of system to 87% for complex websites.
  - Implemented recipe tagging service for raw recipe data comprising the 70+ variety of tags, reducing the dependency on external services and saving the expenditure cost by 92%.

- **Indian Institute of Technology Ropar (IIT-Ropar) [⊕]**   *Jan 2019 - Feb 2021*
  Research Assistant   Ropar, India
  - Worked on vision problems such as image and video super resolution, moving object segmentation, and image restoration (hazy, foggy, lighting clutter ... etc).
  - Developed innovative AI solutions for educational and commercial applications, enhancing learning experiences.

EDUCATION

- **Mohamed bin Zayed University of Artificial Intelligence (MBZUAI)**   *Aug 2023 - Spring 2027*
  PhD, Computer Vision   Abu Dhabi, UAE
  - Thesis: Video Understanding with LMM's, Advisor : Prof. Fahad Khan
  - Research Areas: Multi-Modal Learning, Vision-Language Modeling, Efficient Large Language Models, MultiLingual LLM, LLM Agents, Video Understanding, Image/Video Detection & Segmentation, etc.
  - GPA: 3.74/4.00

- **Shri Guru Gobind Singhji Institute of Engineering and Technology (SGGSIE&T)**   *Jun 2015 - Jul 2019*
  Undergrad, Computer Science   Nanded, India
  - Courseworks Taken: Neural Networks, Machine Learning, Data Structure, Algorithms, Computer Networking, Operating Systems, Cryptography, Distrubuted Systems, Advanced Computer Programming, etc.
  - GPA: 7.7/10.0

PROJECTS

- **Lawa.AI: [Empower your business websites with the power of LLM's]**   *May 2024 - Present*
  Tools: [LLM, Agentic Ecosystem, Knowledge Graph, MultiModal & Document Retrieval]   [⊕]
  - Lawa.AI helps businesses optimize their website & internal organization by providing data-driven insights.
  - By operating directly on the device, lawa.ai ensures data privacy and security for sensitive information.
  - Lawa.ai continually improves through analyzing user interactions to provide more relevant answers over time.
  - Lawa.ai supports multilingual conversations, enabling businesses to reach a global audience.

- **Nutrigenics.Care: [Empowering Health and Wellness for Everyone, Everywhere.]**   *Jan 2024 - Present*
  Tools: [LMM, Recommendation System, MongoDB, Django, Flask]   [⊕]
  - Utilizes advanced AI to instantly track and analyze the nutritional content of meals, ensuring healthier choices.
  - Enables dietitians to track patient meal intake through user-friendly, real-time charts.
  - Engages patients with a user-friendly interface that simplifies the logging process and encourages regular interaction, making it easy to track health data and stay involved in their care.
  - Utilizes advanced AI to provide intelligent recipe suggestions tailored to dietary needs and preferences.

[C.1] Omkar Thawakar, Shravan Venkatraman, Ritesh Thawkar, *et al.* (2025). **EvoLMM: Self-Evolving Large Multimodal Models with Continuous Rewards**. In **CVPR 2026** (Findings)

[C.2] Omkar Thawakar, Dmitry Demidev, *et al.* (2025). **CoVR-R: Reason-Aware Composed Video Retrieval** . In **CVPR 2026** (Findings)

C.3] Dmitry Demidev,..., Omkar Thawakar *et al.* (2025). **Thinking Beyond Labels: Vocabulary-Free Fine-Grained Recognition using Reasoning-Augmented LMMs** . In **CVPR 2026** (Main)

[S.1] A Shaker , ..., Omkar Thawakar, *et al.* (2025). **Mobile-O: Unified Multimodal Understanding and Generation on Mobile Device**. *Under Review*, **Project Page**, **Demo**, **App Store**

[C.4] Omkar Thawakar, Dmitry Demidev, Ritesh Thawkar, *et al.* (2025). **Beyond Simple Edits: Composed Video Retrieval with Dense Modifications**. In **ICCV-2025**

[C.5] Omkar Thawakar, *et al.* (2025). **Llamav-o1: Rethinking step-by-step visual reasoning in llms**. In **ACL-2025** Huggingface Downloads: 50K+

[C.6] Sara Ghaboura,.. Omkar Thawakar, *et al.* (2025). **Time Travel: A Comprehensive Benchmark to Evaluate LMMs on Historical and Cultural Artifacts**. In **ACL-2025**

[C.7] Omkar Thawakar, A Vayani *et al.* (2024). **Mobillama: Towards accurate and lightweight fully transparent GPT**. In **ICLR-2025, SLLM** Spotlight (Top-2%), Huggingface Downloads: 250K+.

[C.8] A Vayani,.. Omkar Thawakar, *et al.* (2024). **All Languages Matter: Evaluating LMMs on Culturally Diverse 100 Languages**. In **CVPR-2025 (Highlight)**

[C.9] Omkar Thawakar, *et al.* (2024). **Composed Video Retrieval via Enriched Context and Discriminative Embeddings**. In **CVPR-2024**

[C.10] Sara Ghaboura, Ahmed Heakl, Omkar Thawakar, *et al.* (2024). **CAMEL-Bench: A Comprehensive Arabic LMM Benchmark**. In **NAACL-2025**

[C.11] Wafa Alghallabi, .. Omkar Thawakar, *et al.* (2024). **Fann or Flop: A Multigenre, Multiera Benchmark for Arabic Poetry Understanding in LLMs**. In *EMNLP-2025 (Main-Track)*

[S.2] Komal Kumar, Tajamul Ashraf, Omkar Thawakar, *et al.* (2024). **LLM Post-Training: A Deep Dive into Reasoning Large Language Models**. In *Arxiv-2025*

[J.1] Omkar Thawakar, *et al.* (2024). **Video instance segmentation in an open-world**. **International Journal of Computer Vision (IJCV) 2024**,DOI:s11263-024-02195-4

[C.12] Omkar Thawakar, *et al.* (2024). **XrayGPT: Chest radiographs summarization using large medical vision-language models**. In **ACL-W-2024**

[C.13] Omkar Thawakar, *et al.* (2023). **3D Mitochondria Instance Segmentation with Spatio-Temporal Transformers**. In **MICCAI-2023**

[C.14] SS Mullappilly, A Shaker, Omkar Thawakar, *et al.* (2023). **Arabic Mini-ClimateGPT: A Climate Change and Sustainability Tailored Arabic LLM**. In **ACL-W-2023**

[C.15] Omkar Thawakar, *et al.* (2023). **Fast Video Instance Segmentation via Recurrent Encoder-Based Transformers**. In *CAIP-2023*

[C.16] Omkar Thawakar, *et al.* (2024). **Video instance segmentation via multi-scale spatio-temporal split attention transformer**. In **ECCV-2022**

[P.1] Omkar Thawakar, *et al.* (2024). **System and method for video instance segmentation via multi-scale spatio-temporal split attention transformer**. **US Patent**, Publication- 2025/12/16, Patent number-12499573.

ENTREPRENEURSHIP EXPERIENCE

• **Cofounder and Acting CEO Lawa.AI**   *July 2024 - Present*
Lawa.AI : Empower your business with the power of Agentic-AI   [🌐]
  ◦ Led end-to-end R&D for cutting-edge AI products, driving innovation across multimodal LLMs, retrieval pipelines, and production-grade Agentic-AI deployments ensuring reliability, accuracy, and measurable impact.
  ◦ Achieved early traction with a $70K seed grant, growing the venture to $130K+ yearly revenue.

• **Cofounder and CTO Nutrigenics.Care**   *Jan 2024 - Present*
Nutrigenics.Care : Empowering Health and Wellness for Everyone, Everywhere.   [🌐]
  ◦ Led end-to-end development of the Nutrigenics platform, including R&D for the Nutrition-GPT AI engine.
  ◦ Collaborated with two healthcare partners to enhance model accuracy, clinical relevance, and platform usability.
  ◦ Raised $100K in grant funding through competitive pitch events.

Skills

- **Programming Languages:** Python, R, Java, C
- **Web Technologies:** Flask, Django, React-JS, MD-Bootstrap, PHP
- **Database Systems:** MongoDB, SQLite, BigQuery
- **Data Science & Machine Learning:** PySpark, Pytorch, Tensorflow, NumPy, Pandas, Matplotlib, Scikit-learn
- **Cloud Technologies:** AWS, Azure, Google Cloud Platform

Honors and Awards

- **Spotlight: ICLR-W2025.** *Feb 2025*
  MobiLlama selected as spotlight research paper (Top 2%) at ICLR-W2025 SLLM (Sparcity in Large Language Models). [⊕]
- **Winner: Khalifa Fund Entrepreneurship Competition** *Dec 2024*
  Secured first place in the Entrepreneurship Competition with 250K AED startup grant support held at Abu Dhabi business week. [⊕]
- **Winner: Sandook-Al-Watan Student Pitch competition** *Nov 2024*
  Secured first place in student pitch competition with 30K AED startup grant support. [⊕]
- **Lawa.AI featured in Panel Discussion at Innovation Xchange in ITU-WTSA 2024** *Nov 2024*
  Represented Lawa.AI as emerging startup from UAE at Innovation Xchange program during WTSA event at New Delhi. [⊕]
- **Microsoft Founders Hub Grant** *July 2024*
  Nutrigenics.Care selected for Microsoft Founders Hub Program with $150K of Technical Support. [⊕]
- **Winner: Pitch Day with Microsoft** *May 2024*
  Nutrigenics.Care secured first place in Pitch Day with Microsoft organized by Incubation & Entrepreneurship Center MBZUAI. [⊕]
- **Winner: Best Project Award for 3D Computer Vision Coursework.** *Nov 2023*
  Worked on 3D Reconstruction using NeRF without using camera parameters for 3D Computer Vision course by Dr. Hao Li. [⊕]
- **Winner: Best Project Award for B.Tech Thesis.** *May 2019*
  Worked on Video Super-Resolution using Recurrent Generative Adversarial Network supervised by Dr. U.V. Kulkarni. [⊕]
- **Winner: Best Reasarch Paper award in International Conference (IC3NS-2018).** *May 2018*
  Worked on the Application of Machine Learning Algorithms for IPM device. [⊕]
- **Winner: Maze Solver Robot Competition for 3 consecutive years 2016-2018 at Pragya SGGS (Tech Event).** *Feb 2018*
  Developed the Maze Solver Robot using Reinforcement Learning that learns the path by itself using Q-Table. [⊕]
- **Winner: Science Exhibition Competition during High School.** *Aug 2012*
  Developed the earthquake detector alarm system using Potentiometer & Relay. [⊕]

Volunteer Experience

- **Teaching Assistant** *Jan 2024 - Dec 2024*
  MBZUAI, Department of Computer Vision [⊕]
  - Assisted professors in preparing and delivering lectures, grading assignments, and maintaining student records.
  - Provided one-on-one and group tutoring sessions for the coursework project and helped the student streamline the project deliverables.

- **Reviewer for Conferences such as CVPR, ICCV, ECCV etc** *July 2022 - Present*
  IEEE, CVF, Springer [⊕]
  - Provided detailed, constructive feedback to authors to improve the quality and clarity of their submissions.
  - The significance, originality, and technical soundness of the proposed methodologies and experimental results were assessed. 25 plus research papers were reviewed from top-tier venues.

References

1. **Prof. Fahad Khan**
   Professor and Deputy Chair Department of Computer Vision
   MBZUAI
   Email: fahad.khan@mbzuai.ac.ae
   Phone: +971 2 811 3254
   Relationship: Primary PhD Advisor
   website | scholar

2. **Dr. Salman Khan**
   Associate Professor, Department of Computer Vision
   MBZUAI
   Email: salman.khan@mbzuai.ac.ae
   Phone: +971 2 811 3256
   Relationship: PhD Co-Advisor
   website | scholar

3. **Prof. Mubarak Shah**
   Trustee Chair Professor, College of Engineering and Computer Science
   University of Central Florida
   Email: shah@crcv.ucf.edu
   Phone: (407) 823-5077
   Relationship: Mentor, Collaborator
   website | scholar

4. **Dr Rao Anwer**
   Assistant Professor, Department of Computer Vision
   MBZUAI
   Email: rao.anwer@mbzuai.ac.ae
   Phone: +971 2 8113253
   Relationship: PhD Co-Advisor
   website | scholar