

YOGESH KUMAR

114, CSE, IIT Jodhpur, India

+91-8630586836 kumar.204@iitj.ac.in

LinkedIn Github Google Scholar yogesh-iitj.github.io

SUMMARY

Ph.D. researcher (Thesis Submitted) specializing in video understanding. Developed state-of-the-art frameworks for spatio-temporal grounding and efficient VideoLLM inference. Strong publication record with 5+ first author papers at top venues including CVPR, ICCV, AAAI, and EMNLP.

EDUCATION

Indian Institute of Technology Jodhpur

M.Tech and PhD in Artificial Intelligence (CGPA: 8.88/10) (Thesis Submitted)

Jodhpur, India

2020 – 2026

- **Thesis:** *Where, How, and When: Query-guided Grounding for Video Understanding*
- **Advisers:** Dr. Anand Mishra
- **Funding:** UGC NET-JRF, Microsoft and Accenture Academic Grants

University of Delhi

B.Sc. (Hons.) in Computer Science (88.7%) and M.Sc. in Computer Science (79.4%)

Delhi, India

2014 – 2019

- **Thesis:** *Extractive Text Summarization*
- **Advisers:** Prof. Vasudha Bhatnagar

EXPERIENCE

IIT Jodhpur

Senior Project Scientist

Jodhpur, India

Jan/2026 – Present

- Worked on Video to Video Retrieval, the work is under submission in ECCV'26

Microsoft Academic Partnership (MAPG)

Project lead

Jodhpur, India

Jul/2024 – Jun/2025

- Worked with Microsoft Bing Hyderabad on long-video action search using text/video queries; published at ICCV'25 and ICVGIP'24.

Accenture Academic Collaboration

Project lead

Jodhpur, India

Jul/2023 – Jun/2024

- Worked with Accenture Bangalore on Localizing unseen objects and their interaction in videos; published at CVPR'23 and AAAI'24.

PUBLICATIONS

1. **Yogesh Kumar**, Anand Mishra, *Few-Shot Referring Relationships in Videos*, IEEE/CVF Conference on Computer Vision and Pattern Recognition (**CVPR**) 2023 [paper][poster][project][code].
2. **Yogesh Kumar**, Saswat Mallick, Anand Mishra, Sowmya Rasipuram, Anutosh Maitra, Roshni Ramnani, *QDETRv: Query-Guided DETR for One-Shot Object Localization in Videos*, Association for the Advancement of Artificial Intelligence (**AAAI**) 2024 [paper][poster][code].
3. **Yogesh Kumar**, Uday Agarwal, Manish Gupta, Anand Mishra, *Aligning Moments in Time using Video Queries*, IEEE/CVF International Conference on Computer Vision (**ICCV**) 2025 [paper][code][dataset].
4. **Yogesh Kumar**, *Language-Guided Temporal Token Pruning for Efficient VideoLLM Processing*, Empirical Methods in Natural Language Processing (**EMNLP Main**) 2025 [paper][code]
5. **Yogesh Kumar**, Anand Mishra, *Temporal Object-Aware Vision Transformer for Few-Shot Video Object Detection*, Association for the Advancement of Artificial Intelligence (**AAAI**), 2026 [paper][code]
6. **Yogesh Kumar**, Uday Agarwal, Manish Gupta, Anand Mishra, *Moment Alignment Transformer for Moment Retrieval using Video Queries*, IEEE Transactions on Pattern Analysis and Machine Intelligence (**T-PAMI**), Under Review. [paper]
7. Shreya Shukla, Prajwal Gatti, **Yogesh Kumar**, Vikas Yadav, Anand Mishra, *Towards Making Flowchart Images Machine Interpretable*, International Conference on Document Analysis and Recognition (**ICDAR**) 2023 [paper][poster][project][code].
8. Uday Agarwal, **Yogesh Kumar**, Abu Shahid, Prajwal Gatti, Manish Gupta, Anand Mishra, *Chapter-Based Video Moment Retrieval using Natural Language Queries*, Indian Conference on Computer Vision, Graphics and Image Processing (**ICVGIP**) 2024 [code and dataset].

9. **Yogesh Kumar**, *VideoLLM Benchmarks and Evaluation: A Survey*, (**arXiv**) 2025 [paper]
10. **Yogesh Kumar**, *Cross-Modal Graph Reasoning for Training-Free Spatio-Temporal Video Grounding*, (**Under Submission, ECCV 2026**).
11. Uday Agarwal, Alik Sarkar, **Yogesh Kumar**, Manish Gupta, Anand Mishra, *When “Video” is its Own Best “Query”*, (**Under Submission, ECCV 2026**).

AWARDS AND FELLOWSHIPS

- BMVC 2024 and 2025 Outstanding Reviewer
- UGC National Eligibility Test-Junior Research Fellowship (NET-JRF) in CS 2020 (Percentile 99.96, Rank 17)
- Graduate Aptitude Test in Engineering (GATE) in CS 2020 (Rank 605)
- Microsoft Academic Partnership Grant (MAPG) 2024, *Principal Investigator: Dr Anand Mishra*
- Accenture Gift Grant 2023, *Principal Investigator: Dr Anand Mishra*
- ACM/IARCS Travel Grants for Attending CVPR 23 (Canada, Vancouver) and AAI 26 (Singapore)
- Google, Microsoft Travel Grant for attending CVPR 2023
- Selected to Attend Research Week with Google in 2022, 2024, and 2025
- Selected to Attend Microsoft Research India Academic Summit 2025

MENTORING AND LEADERSHIP

Research Mentorship (Advised by Dr. Anand Mishra)

- Mentored **6+ students** (B.Tech and M.Tech) in Computer Vision and NLP, resulting in **3 peer-reviewed publications** at top-tier venues (ICCV, ICDAR, ICVGIP) and **3 successfully defended M.Tech theses**.
- **Shreya Shukla** (B.Tech): Guided research on flowchart to code generation, leading to a publication at **ICDAR 2023**.
- **M.Tech Theses**: Supervised graduate research for Apoorv Shekar (Predicate Grounding), Kranti Prakash (Video QA), and Vishal Dangiwalla (Grounded QA in Videos).

Teaching Assistantships, IIT Jodhpur

- Assisted in **7+ courses** (including Pattern Recognition, Machine Learning, and Computer Vision) for classes of 100+ students. Conducted lab sessions, designed assignments, and evaluated technical projects.
- **Tutorial**: Delivered a session on *Seq2seq models* at the Winter School on Responsible AI, 2022.

ACADEMIC SERVICES

Reviewer

- **Conferences**: ICLR 2024, 2025; BMVC 2024 & 2025 (Outstanding Reviewer); ECCV 2024, 2026; NeurIPS 2023, 2025; CVPR 2026.
- **Journals**: TPAMI, TMM, TIST, TIP

Volunteer

- AAI 2026 (Singapore); EMNLP 2025 (Suzhou, China); Winter Schools 2022 & 2024, IIT Jodhpur; NCVPRIPG 2023, IIT Jodhpur.

TALKS AND PRESENTATIONS

- Presented research on video moment retrieval and one-shot object localization at Vision India (ICVGIP 2024–25), with publications at ICCV 2025 and AAI 2024.
- Showcased work on visual relationship grounding in videos at Vision India (ICVGIP 2023), AIML-Sys 2023, and ACM-SIGKDD Workshop 2023, tied to CVPR 2023.

TECHNICAL SKILLS

Languages: Python, C, C++ **Libraries:** PyTorch, TensorFlow, OpenCV, Pandas, Numpy, Scikit-learn **Tools:** Jupyter/IPython Notebook, Git, LaTeX