

Anil Batra

✉ anilbatra2185@gmail.com

🐦 @_anilbatra

in anil-batra

🌐 <https://anilbatra2185.github.io/>

Research Interests

- Multimodal AI ♦ *Long-Context Modeling, Memory & Reasoning, Conditional Text Generation*
- Ego/Exo-centric Video ♦ *Human-Object Understanding, 2D/3D Spatial-Temporal Scene Graph*
- Other ♦ *Low-resource & Light-weight Models, Generative Modeling*

Education

- 2020 – 2025 ♦ **Ph.D. Computer Science, University of Edinburgh, UK.**
Supervisors: [Prof. Frank Keller](#), [Prof. Marcus Rohrbach](#), [Dr. Laura Sevilla](#)
Research Focus: *Understanding the structural aspects of long Exo-centric videos to enhance multimodal learning, aiming to develop efficient models that can plan, reason, and execute goal-oriented tasks.*
Expected Graduation: *September 2025*
- 2016 – 2019 ♦ **M.Sc. Computer Science, IIIT Hyderabad, India**
Supervisors: [Prof. C.V. Jawahar](#)
Thesis: *Road Topology Extraction from Satellite images by Knowledge Sharing.*
GPA: 8.44/10
- 2003 – 2007 ♦ **B.Tech. Electronics & Communication, Punjab Technical University, India.**
GPA: 8.38/10

Employment

- 2022/06 – 2022/09 ♦ **Research Intern, Mediatech Research, UK**
Research Focus: Focused on diffusion model baselines for unconditional image generation, laying the groundwork for future work '[Image generation with shortest path diffusion](#)', accepted at ICML 2023.
Manager: Alberto Bernacchia
- 2020/05 – 2020/09 ♦ **Sr. Data Scientist, ESRI, India**
Research Focus: Incorporated my published road extraction work [arc-gis](#) library for future use by OSM community engineers.
Manager: Rohit Singh
- 2019/01 – 2020/04 ♦ **Research Engineer, Pro @ Facebook, India**
Research Focus: Train and deploy road extraction pipeline to OSM maps. [Technical Blog by Meta](#)
Manager: Saikat Basu

Research Publications

Conference Proceedings

- 1 **Anil, Batra**, L. Sevilla-Lara, M. Rohrbach, and F. Keller, "Predicting implicit arguments in procedural video instructions," in *ACL (Main)*, 2025. 🔗 URL: <https://arxiv.org/pdf/2505.21068>.
- 2 D. Gautier, L. Olga, and **Anil, Batra**, "Cast: Cross-modal alignment similarity test for vision language models," in *COLING*, 2025. 🔗 URL: <https://aclanthology.org/2025.coling-main.93.pdf>.

- 3 B. Meinardus, H. G. Rodriguez, A. Batra, A. Rohrbach, and M. Rohrbach, "Chrono: A simple blueprint for representing time in mllms," in *ICCV, Workshop on Multimodal Representation and Retrieval*, 2025. [URL: https://arxiv.org/pdf/2406.18113](https://arxiv.org/pdf/2406.18113).
- 4 **Anil, Batra**, D. Moltisanti, L. Sevilla-Lara, M. Rohrbach, and F. Keller, "Efficient pre-training for localized instruction generation of videos," in *European conference on computer vision*, Springer, 2024. [URL: https://www.ecva.net/papers/eccv_2024/papers_ECCV/papers/05611.pdf](https://www.ecva.net/papers/eccv_2024/papers_ECCV/papers/05611.pdf).
- 5 A. Das, S. Fotiadis, **Anil, Batra**, et al., "Image generation with shortest path diffusion," in *International Conference on Machine Learning*, PMLR, 2023, pp. 7009–7024. [URL: https://arxiv.org/pdf/2306.00501](https://arxiv.org/pdf/2306.00501).
- 6 **Anil, Batra**, S. N. Gowda, F. Keller, and L. Sevilla-Lara, "A closer look at temporal ordering in the segmentation of instructional videos," in *British Machine Vision Conference*, 2022. [URL: https://arxiv.org/pdf/2209.15501](https://arxiv.org/pdf/2209.15501).
- 7 **Anil, Batra**, S. Singh, G. Pang, S. Basu, C. Jawahar, and M. Paluri, "Improved road connectivity by joint learning of orientation and segmentation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 10 385–10 393. [URL: https://openaccess.thecvf.com/content_CVPR_2019/papers/Batra_Improved_Road_Connectivity_by_Joint_Learning_of_Orientation_and_Segmentation_CVPR_2019_paper.pdf](https://openaccess.thecvf.com/content_CVPR_2019/papers/Batra_Improved_Road_Connectivity_by_Joint_Learning_of_Orientation_and_Segmentation_CVPR_2019_paper.pdf).
- 8 S. Singh, **Anil, Batra**, L. Torresani, et al., "Self-supervised feature learning for semantic segmentation of overhead imagery," BMVA Press, 2018. [URL: http://bmvc2018.org/contents/papers/0345.pdf](http://bmvc2018.org/contents/papers/0345.pdf).

Skills

Coding	◇ Python (advanced), C++ (beginner), C#.Net (intermediate)
ML Frameworks	◇ PyTorch (advanced), Jax (beginner), TensorFlow (beginner)
Misc.	◇ GDAL (beginner), SQL (beginner), Matlab (beginner)
Languages	◇ English (advanced), Hindi (native), Punjabi (native)

Miscellaneous Experience

Honors & Awards

2024	◇ Outstanding reviewer award at BMVC, Glasgow, UK
2020-2024	◇ Ph.D. Funding, UKRI CDT in NLP, University of Edinburgh, Scotland

Teaching

2021 - 2024	◇ Teaching Support for Applied Machine Learning. Roles: Marker, Exam Marker, Tutor
2024	◇ Doing Research in NLP course for first-year PhD students. Roles: Marker
2023	◇ Teaching Support for Natural Language Generation & Understanding. Roles: Marker, Tutor
2021	◇ Teaching Support for Data Mining and Exploration. Roles: Demonstrator, Tutor

Reviewer

2025	◇ I have contributed as a reviewer to ACL-ARR.
2024 +	◇ I have contributed as a reviewer to AAAI.

Miscellaneous Experience (continued)

2021 + ◇ I have contributed as a reviewer to CVPR/ECCV/ICCV/BMVC.