Yash Sanjay Bhalgat

Senior Research Scientist, Google DeepMind

EMAIL: *yashbhalgat95@gmail.com* / *yashbhalgat@google.com*
WEBPAGE: *yashbhalgat.github.io* ◇ LINKEDIN: *yashbhalgat*

WORK EXPERIENCE

Google DeepMind, *Senior Research Scientist – Gemini Omni (Veo)* [Nov '25 - Present]

- Core contributor in the "pretraining" and "video editing" pillars of the Gemini Omni video generation model [**Blog**].
- **Lead captain** for Omni v1.5 – built the code infrastructure for Torsopatch pretraining, ran several model training ladders to incorporate critical capabilities into the model, and launched the final model training run for Omni v1.5.
- Co-led 3D/camera capabilities in Omni v1 – built a camera-captioning pipeline with significantly stronger text-to-camera control than Veo 3.1; also, shipped novel "Camera Reshoot" and "Camera Motion Copy" capabilities.
- Built the "Variable-Length Captioning" pipeline for video editing, enabling segment-level granularity for edits in Omni.
- Received 3 peer bonuses across 3 different projects – rarely given to recognize exceptional work.

Meta Reality Labs, *Research Scientist Intern* [Apr '25 - Sep '25]

- Built 4D motion-conditioned video generation world models for robotics.

Multiple startups, *Part-time AI Consultant* [Feb '23 - Mar '25]

- *SoftEye Inc (acquired by TDK Corp)*: Developing real-time 3D vision algorithms for augmented reality on smart glasses.
- *GoBubble Ltd*: Finetuned and deployed Large Language Model (LLM) models to moderate multimodal data online.
- *Togal.AI*: Building multimodal solutions for detecting and understanding project features on architectural plans/drawings.

Qualcomm AI Research | *Senior Machine Learning Researcher* [Nov '20 - Jul '21]
Machine Learning Researcher [Jun '19 - Oct '20]

- Spearheaded the ultra-low resource always-on vision project from model design, quantization to final hardware mapping
- Filed 12 inventions in 2020-21 of which 6 ideas have been filed for patent protection. Notable works on 3D hand-pose estimation [DIR-Net], low-bit quantization [LSQ+, QKD], structured [StructConv] and unstructured [LTP] pruning
- Led Qualcomm's team in the MicroNet Challenge at NeurIPS 2019, and achieved 3rd position in ImageNet track [Code]
- Managed/mentored interns - Jangho Kim and John Yang (PhD @ SNU) with contributions to the AR/VR project

Voxel51, Inc., *Computer Vision & Machine Learning Engineer* [Feb '19 - May '19]

- Researched and developed production pipelines for real-time vehicle tracking for querying on large-scale video databases

EDUCATION

University of Oxford **Oct '21 - Nov '25**
PhD (DPhil), Computer Vision and Machine Learning @ Visual Geometry Group (VGG)
Thesis: *From video to virtual: object-centric 3D scene understanding from videos*
Advisors: Prof. Andrew Zisserman, Prof. Andrea Vedaldi, Dr. João Henriques, Dr. Iro Laina

University of Michigan, Ann Arbor **Sep '17 - Dec '18**
Masters, Computer Science and Engineering

Indian Institute of Technology, Bombay **Jul '13 - May '17**
B.Tech. (with *Honors*) in Electrical Engineering with *Minor* in Computer Science
Thesis: *Scattering Wavelets Network based Latent Fingerprint Enhancement*, **Advisor**: Prof. Vikram Gadre
Undergraduate Research Award (URA 02) for exceptional Bachelors Thesis

SELECTED PUBLICATIONS

Conference Publications (Full list: *Google scholar*) * equal contribution

12. Do 3D Large Language Models Really Understand 3D Spatial Relationships?. [Paper] [Project]
ICLR, 2026 (**Best paper honorable award, 3D-LLM/VLA Workshop, CVPR 2026**). X. Ma, T. Sun, S. Chen, *Yash Bhalgat*, J. Gu, A. Chang, I. Armeni, I. Laina, S. Peng, V. Prisacariu.
11. Jamais Vu: Exposing the Generalization Gap in Supervised Semantic Correspondence. [Paper] [Code]
NeurIPS, 2025. Octave Mariotti, Zhipeng Du, *Yash Bhalgat*, Oisin Mac Aodha, Hakan Bilen.
10. Reflecting Reality: Enabling Diffusion Models to Produce Faithful Mirror Reflections. [Paper] [Code] [Dataset]
3DV, 2025. A. Dhiman*, M. Shah*, R. Parihar, *Yash Bhalgat*, L. Boregowda, R Venkatesh Babu.
9. GS-CPR: Efficient Camera Pose Refinement via 3D Gaussian Splatting. [Paper] [Code]
ICLR, 2025. C. Liu, S. Chen, *Yash Bhalgat*, S. Hu, M. Cheng, Z. Wang, V. Prisacariu, T. Braud.

8. 3D-Aware Instance Segmentation and Tracking in Egocentric Videos. [Paper]
   ACCV, 2024. *Yash Bhalgat**, Vadim Tschernezki*, Iro Laina, João Henriques, Andrea Vedaldi, Andrew Zisserman.
7. N2F2: Hierarchical Scene Understanding with Nested Neural Feature Fields. [Paper]
   ECCV, 2024. *Yash Bhalgat*, Iro Laina, João Henriques, Andrew Zisserman, Andrea Vedaldi.
6. SiLVR: Scalable Lidar-Visual Reconstruction with Neural Radiance Fields for Robotic Inspection. [Paper]
   ICRA, 2024. Yifu Tao, *Yash Bhalgat*, Lanke Frank Tarimo Fu, Matias Mattamala, Nived Chebrolu, Maurice Fallon.
5. Neural Refinement for Absolute Pose Regression with Feature Synthesis. [Paper]
   CVPR, 2024. Shuai Chen, *Yash Bhalgat*, Xinghui Li, Jiawang Bian, Kejie Li, Zirui Wang, Victor Adrian Prisacariu.
4. Contrastive Lift: 3D Object Instance Segmentation by Slow-Fast Contrastive Fusion. [Paper][Code]
   NeurIPS, 2023 (**Spotlight**). *Yash Bhalgat*, Iro Laina, João Henriques, Andrea Vedaldi, Andrew Zisserman.
3. A Light Touch Approach to Teaching Transformers Multi-view Geometry. [Paper]
   CVPR, 2023. *Yash Bhalgat*, João Henriques, Andrew Zisserman.
2. A Prompt Array Keeps the Bias Away: Debiasing Vision-Language Models with Adversarial Learning. [Paper]
   AACL-IJCNLP, 2022. Hugo Berg, Siobhan Hall, *Yash Bhalgat*, Wonsuk Yang, Hannah Kirk, A. Shtedritski, M. Bain.
1. Structured Convolutions for Efficient Neural Network Design. [Paper]
   NeurIPS, 2020. *Yash Bhalgat*, Yizhe Zhang, Jamie Lin, Fatih Porikli.

SELECTED INTERNSHIPS & PROJECTS

[Project] NeurIPS '19 MicroNet challenge - 3rd place, ImageNet track [Code] [Jul '19 - Oct '19]
- Designed fast evolutionary mixed precision quantization: 8x compression EfficientNet-B0/MixNet-S, <1% accuracy loss

[Internship] IBM Almaden Research Center, *Mentor - Zhe Liu, Pritam Gundecha* [Summer '18]
- Proposed teacher-student learning paradigm for task-agnostic classification in presence of label noise in train data [Paper]

[Internship] IFP Energies nouvelles, Paris, *Mentor - Laurent Duval* [Summer '17]
- Real-time sparsity-based deformation-invariant feature extraction for large geophysical images (ICASSP 2018) [Paper]

[Internship] IBM Research, Bangalore, *Mentor - Vikas Raykar* [Summer '16]
- Joint multi-modal representations for e-commerce catalog search by visual attributes *without* manual tagging

SKILLS

Languages	Python (proficient), C++ (moderate), Julia, MATLAB, Verilog, Bash, LaTeX
Frameworks	PyTorch, Jax (proficient), CUDA, LangChain, DSPy, 🤗 Diffusers, TensorFlow & Keras (basic), git, slurm

TEACHING EXPERIENCE

University of Oxford, *Teaching Assistant*	Computer Vision, *w/ Profs Andrea Vedaldi, Andrew Zisserman*	[Hillary '22, '24, '25]
	Computer Graphics, *with Dr. Jassim Happa, Stuart Golodetz*	[Hillary '22]
	Artificial Intelligence, *with Prof. Bernardo Cuenca Grau*	[Hillary '22]
University of Michigan, *Graduate Student Instructor*	Computational Data Science, *with Prof. Raj Nadakuditi*	[Fall '18]
	Introduction to Logic Design, *with Prof. Matthew Smith*	[Winter '18]
IIT Bombay, *Teaching Assistant*	Wavelets, *with Prof. Vikram Gadre*	[Fall '16, Winter '17]
	Quantum Mechanics and Applications, *with Prof. Siva Prasad*	[Fall '14, Winter '15]

SCHOLASTIC ACHIEVEMENTS

- Undergraduate Research Award (URA 02) for exceptional work during Bachelors Thesis at IIT Bombay
- Cargill Global Scholarship 2014-15 and 2015-16 for excellence in leadership and academic skills
- All India Rank **12** in IITJEE-Mains exam among 1 million candidates
- All India Rank **155** in IITJEE-Advanced exam among 150,000 candidates
- All India Rank **60** in KVPY Scholarship exam by Govt. of India among 0.2 million candidates
- Selected among National Top 30 (for OCSC camp) for International Astronomy Olympiad '13
- Selected among top 300 participants of India to compete in **all three national olympiads**: INPhO (Indian National Physics Olympiad), INChO (Chemistry), INAO (Astronomy)
- Visharad Degree (*equivalent to* Bachelors in Music) in Indian Classical Music for playing Tabla