Sauradip Nag

Researcher and Engineer

+1 7783234784 sauradipnag95@gmail.com Vancouver, CA
in/sauradip github.com/sauradip x.com/saura

Career Highlights

- **14 Co-authored A^* publications (w/ 9 First-author):** 3 papers in ECCV, 4 papers in SIGGRAPH, 2 papers in AAAI (1 Oral) and 1 paper each in CVPR, ICCV, NeurIPS, ICLR, ICASSP covering Vision, Graphics and NLP domain.
- **More than 4 years of Engineering Experience:** Research Engineer on *Deep Learning for Embedded Systems* for 2 years + *Software Development and Research Scientist/Engineering Internship* experience for 2.5 years
- **4 years of Teaching and Supervising Experience:** Supervised 6 PhD students combined in University of Surrey, Simon Fraser University for 4 years + Teaching Assistant for 2 courses for 8 months in University of Surrey.

Area of Expertise

a) Large Scale Training (DDP,MP,PEFT) - b) Foundation Models (LLM, Stable Diffusion, SAM, LLM, CLIP) - c) Generative Models (Flow-Matching, Diffusion, GAN, VAE) - d) Programming Language (Python, C++) - e) Deep Learning Stack (PyTorch, Tensorflow) - f) 3D Engine (Blender, Maya)

Professional Experience

Postdoctorate Researcher, *(Simon Fraser University)* **Burnaby, Canada** *03/2024 - present*
- Published 6 top conference paper in TOG, SIGGRAPH, SIGGRAPH ASIA, ICLR, NeurIPS and supervising 4 PhD students.
- Developing foundation models for 4D generation in a feedforward fashion which can generate results in 10 seconds.
- Fine-tuned a custom 3D and Video segmentation model which increases efficiency by 3x and better quality.
- Wrote proposals to obtain research grant worth 500K USD.

Research Scientist Intern, *(Huawei Noah's Ark Lab)* **London, UK** *11/2022 - 08/2023*
- Published 2 top conference papers in ICCV and AAAI (Oral).
- Trained a Large Scale Diffusion Model for Video and Audio Understanding.
- Worked on 3D Human Mesh Texturing and 3D Diffusion Models.

Research Associate/Engineer, *(ICSR, IIT Madras)* **Chennai, India** *08/2018 - 03/2020*
- Developed the first night-time scene segmentation model which can run at 15 FPS, published as Spotlight Paper in ICIP.
- Implemented RGBD-SLAM for edge devices with a RoboticOS wrapper for realtime path planning and improved accuracy by 30%
- Supervised a team of 2 Masters and 1 PhD student for succesful project delivery ahead of time and improved efficiency.

Computer Vision Developer Intern, *(GISCLE Systems)* **Bengaluru, India** *01/2018 - 08/2018*
- Worked on Advance Lane Detection System for Autonomous Cars in India
- Solved $0 \rightarrow 1$ problem in Level2 Autonomous Driving using computer vision.
- Collected and Annotated 60+ Hours of Driving Videos and trained Deep Learning Models.

Software Development Engineer Intern, *(Apnastudy Technologies)* **Texas, USA** *02/2016 - 7/2017*
- Worked on developing AI based Learning Management Systems using MEAN Stack
- Reached 110% of customer acquisition target, capturing 1000 new subscribers in first month of launch.
- Managed both Frontend and Backend Production Servers

Education

PhD Electrical Engineering *University of Surrey* **Guildford, UK** *2020-2024*
Research Focus on Generative & Multi-modal Video Understanding; Published 6 First Author paper in Top Conferences liek CVPR,ICCV,ECCV,AAAI ; Teaching Assistant in 2 CS courses; Research Intern for 8 months at Huawei London.

B.Tech Computer Engineering *Kalyani Government Engineering College* **West Bengal, India** *2014-2018*
Research Focus on Vector Graphics and Image Understanding; Published paper in Top Conferences like ICASSP, ICDAR and journals like Pattern Recognition; Interned in 2 startups for combined 2 years of Software Development experience

Selected Publications*

- **CountLoop: Training-Free High Instance Image Generation via Agent Guidance** *Under Review* *2025*
  Anindya Mondal, **Sauradip Nag***, Ayan Banerjee*, Xiatian Zhu, Josep Lados, Anjan Dutta
  We propose the first training-free model for high instance image generation using MLLM as a critique

- **Articulate That Object Part: 3D Part Articulation from Text and Motion** *Trans. of Graphics (TOG)* *2026*
  Aditya Vora, **Sauradip Nag**, Kai Wang, Richard Zhang
  We propose a novel approach to articulate 3D mesh fine-tuned using personalized Multi-View T2V priors

- **RespoDiff: Dual-Module Bottleneck Transformation for Safe Image Generation** *NeurIPS* *2025*
  Silpa VS, **Sauradip Nag**, Muhammad Rana, Serge Belongie, Anjan Dutta
  We solve the problem of Responsible T2I generation using a new score distillation objective

- **ASIA: Adaptive 3D Segmentation using Few Image Annotations** *SIGGRAPH Asia* *2025*
  Sairaj Kishore, **Sauradip Nag***, Aditya Vora*, Ali-Mahdavi Amiri, Richard Zhang
  We propose a customized 3D mesh segmentation approach using T2I Diffusion with Multi-View consistency

- **In-2-4D: Inbetweening from Two Single-View Images to 4D Generation** *SIGGRAPH Asia* *2025*
  Sauradip Nag, Daniel Cohen-Or, Richard Zhang, Ali-Mahdavi Amiri
  We propose a new problem of 4D generation from sparse views in a novel divide-and-conquer strategy

- **Cora: Correspondence-aware image editing using few step diffusion** *SIGGRAPH* *2025*
  Amir M.*, Aryan Mikaeili*, **Sauradip Nag**, Negar Hassanpour, Andrea Tagliasachi, Ali-Mahdavi Amiri
  We propose the first few-step non-rigid image editing approach using point correspondence

- **SMITE: Segment Me in Time** *ICLR* *2025*
  Amir Mohammadi, **Sauradip Nag**, Saeid Asgari Taghanaki, Andrea Tagliasacchi, Ali-Mahdavi Amiri
  We solved a novel task of few-shot customized video segmentation using T2V Diffusion

- **DiffTAD: Temporal Action Detection with Proposal Denoising Diffusion** *ICCV* *2023*
  Sauradip Nag, Xiatian Zhu, Jiankang Deng, Yi-Zhe Song, Tao Xiang
  We proposed and trained the first Transformer based Diffusion model for Video Action Understanding

- **Multi-Modal Few-Shot Temporal Action Detection** *Arxiv Preprint* *2023*
  Sauradip Nag, Mengmeng Xu, Juan Perez Rua, Xiatian Zhu, Bernard Ghanem, Yi-Zhe Song, Tao Xiang
  We trained a large scale VLM model for Videos end-to-end to solve a new task of Multi-Modal Few-Shot Learning

- **Post-processing Temporal Action Detection** *CVPR* *2022*
  Sauradip Nag, Xiatian Zhu, Yi-Zhe Song, Tao Xiang
  We proposed a training-free method to solve Video Action Localization using Taylor Series Expansion

- **Zero-Shot Temporal Action Detection via Vision-Language Prompting** *ECCV* *2022*
  Sauradip Nag, Xiatian Zhu, Yi-Zhe Song, Tao Xiang
  We proposed the first VLM based Video Action Localization method

- **Semi-Supervised Temporal Action Detection with Proposal-Free Mask Learning** *ECCV* *2022*
  Sauradip Nag, Xiatian Zhu, Yi-Zhe Song, Tao Xiang
  We proposed a SSL based Video Action Localization method to reduce data bias

- **Proposal-Free Temporal Action Detection using Global Segmentation Mask Learning** *ECCV* *2022*
  Sauradip Nag, Xiatian Zhu, Yi-Zhe Song, Tao Xiang
  We proposed a new single-stage Video Localization framework which can be trained under 15 mins

Honors and Awards

- **AAAI Top 2% Paper** Oral in AAAI 2024 for *DiffSED* paper
- **ICIP Top 10% Paper:** Spotlight in ICIP 2019 for *What's There in the Dark* paper
- **Runners Up CVPR Workshop:** Award in FGVC9 Worskhop in 2022 for *Fine-grained Retrieval* work
- **IflyTek Fully Funded PhD Scholarship:** Received fully funded scholarship for 3.5 Years of PhD in UK
- **Oxford ML Summer School:** Presented a invited talk on *Video Understanding in Industry*
- **Smart India Hackathon Silver:** Award in 2019 Hackathon on solving practical computer vision problem