Navlika Singh

LinkedIn Github Google Scholar

Email: ns1324@ic.ac.uk
Mobile: +44-07587292943

Professional Summary

ML Researcher experienced in designing, training, and evaluating deep learning architectures across multimodal systems - audio-visual, speech, and vision-language. Five academic papers, reviewer for CVPR/ECCV/NeurIPS, and hands-on experience building LLM-based pipelines, evaluation frameworks, and production AI systems at speed. Passionate about efficient architectures, model scaling, and working at the intersection of research and real-world constraints.

Education

- **Imperial College London** — London, United Kingdom
  MSc in Computing (Artificial Intelligence & Machine Learning) | Term 1: Distinction — *Sep '25 - Sep '26*

  Relevant Coursework: Statistical Information Theory; Machine Learning Systems and Hardware; Mathematics for Machine Learning; Introduction to Machine Learning

- **Indian Institute of Technology (IIT)** — Jodhpur, India
  B.Tech. in Artificial Intelligence & Data Science | Examination & Thesis: Distinction — *Dec '20 - May '24*

  Relevant Coursework: Probability, Statistics & Stochastic Processes; Pattern Recognition & ML; Human–Machine Interaction; Deep Learning; Dependable AI; Advanced ML; Speech Understanding.

Publications

- **SafeCoVT: A Plug-and-Play Framework for Safety-Critical Applications**: (paper)
  Navlika Singh, Oana-Maria Camburu
  2025. In **Internal Conference** at Imperial College London
- **BLADE: Bias-Linked Adaptive DEbiasing**: (Under Review - paper)
  Piyush Arora*, **Navlika Singh***, Vasubhya Diwan*, Pratik Mazumder*
  2025. In **Pattern Recognition** Journal
- **When Big Models Train Small Ones: Label-Free Model Parity Alignment for Efficient Visual Question Answering using Small VLMs**: (Accepted - paper)
  Abhirama Subramanyam Penamakuri*, **Navlika Singh***, Piyush Arora*, Anand Mishra
  2025. In **EMNLP Main** Track
- **Dataset Selection is Critical for Effective Pre-Training of Fish Detection Models for Underwater Video**: (Accepted - paper)
  Devi Ayyagari, Talukder Wasi Alavi, **Navlika Singh**, Joshua Barnes, Corey Morris, Christopher Whidden
  2025. In **ICES Journal of Marine Science**
- **UniPreCIS: A data pre-processing solution for collocated services on shared IoT**: (Accepted - paper)
  Anirban Das, **Navlika Singh**, Suchetana Chakraborty
  2024. In Journal of **Future Generation Computer Systems**

*Equal contribution.

Work Experience

- **Metafusion** — Noida, India
  Founding Engineer - Computer Vision Department — *May '24 - Aug '25*
  - **Lightweight VLMs for Surveillance**: Built an end-to-end pipeline for analyzing vehicles and pedestrians in CCTV footage. Localized subjects with YOLOv11, and used lightweight VLMs (SmolVLM-256M, Florence-2-base) to generate attribute labels and natural language summaries in real time (under 1 sec).
  - **Safe-City Applications**: Optimized VLMs (Qwen2.5-VL-3B) for low-resource city surveillance, achieving 90% F1 score on action-based applications.
  - **Intelligent Traffic Management Systems (ITMS) Applications**: Built a helmet and triple-riding violation detection system, achieving +12% overall and +17.5% accuracy improvement for triple-riding detection. Also, re-designed License Plate Detection (LPD) from a bounding box to a keypoint-based approach.

  Keywords: Real-Time Inference, Model Optimisation, Prompt Engineering, Lightweight VLMs, Object Detection, Vision-Language Models, Low-Resource Deployment, Edge AI, Production ML, Kubernetes, VLLM, GCP

Research Experience

- **Imperial College London** — London, United Kingdom
  Student Researcher - iBUG Labs (Head: Prof. Maja Pantic FRS) — *Oct '25 - Present*
  - **Matryoshka Speech Tokenization** (*Dr. Umberto Campebella*, *Dr. Stavros Petridis*)
    - Investigating Matryoshka Representation Learning for speech tokenizers to support dynamic frame-rate and positive cross-granularity transfer, improving low-bitrate performance while reducing training cost.

  Keywords: Speech Tokenisation, Matryoshka Representation Learning, Efficient Inference, Low-Bitrate Optimisation, Self-Supervised Learning, PyTorch, Slurm

- **Lip-Sync Personalization** (*Dr. Antoni Bigata*, *Dr. Stavros Petridis*)
  - Developed a LoRA-based framework for cross-identity speaking style transfer in audio-driven lip synchronization, with a three-stage regularization pipeline to prevent identity leakage from few-shot training.
  - Curated an audio-visual dataset of speaker-specific speaking styles and introduced a dedicated metric for speaking style transfer fidelity, replacing generic lip-sync scores or human evaluation.
  - Achieved SOTA speaking style transfer while maintaining lip-sync and image quality, validated as a plug-and-play solution across multiple lip-sync backbones.
  - This work is currently under-review at **Interspeech 2026**.

  Keywords: LoRA, Parameter-Efficient Fine-Tuning, Multimodal Learning, Audio-Visual Generation, Few-Shot Learning, Generative Modelling, PyTorch
- **Audio-Visual Speech Recognition** (*Dr. Umberto Campebella*, *Dr. Stavros Petridis*)
  - Analyzed how acoustic noise propagates through LLM layers in an LLM-based AVSR model via layer-wise cosine similarity, finding that noisy audio tokens diverge early and corrupt visual representations.
  - Proposed a lightweight method integrating variational information bottleneck layers into the LLM backbone to regularize representations, reducing the clean-noisy representational gap.
  - Evaluated across multiple noise conditions, demonstrating improved noise robustness while preserving strong performance on standard AVSR benchmarks.
  - This work is currently under-review at **Interspeech 2026**.

  Keywords: Audio-Visual Speech Recognition, Multimodal Fusion, Noise Robustness, Variational Information Bottleneck, LLM Interpretability, Efficient Fine-Tuning, PyTorch

- **Imperial College London** — London, United Kingdom
  Student Researcher, Supervisor: Prof. Oana-Maria Camburu — *Nov '25 - Present*
  - **Latent Reasoning Models** (MSc Thesis)
    - Investigating the safety, interpretability, and monitorability of latent reasoning models.
  - **Safe Chain-of-Thought Visual Reasoning** (Independent Study Option)
    - Conducted a systematic study of CoT reasoning variants on safety-critical tasks, identifying key failure modes in text-only CoT including visual hallucination, incomplete coverage, and constraint misinterpretation.
    - Proposed SafeCoVT: a plug-and-play framework enabling MLLMs to autonomously orchestrate visual perception tools through iterative reasoning, without domain-specific fine-tuning.
    - Developed the Comprehensive Safety Score (CSS) for holistic safety evaluation; SafeCoVT up to 55.57% improvement in constraint violation rate and 47.95% in CSS over direct question-answering on the ASIMOV-2.0 benchmark.
    - This work was presented at an **internal conference** at Imperial College London.

  Keywords: AI Safety, Interpretability, Monitorability, Chain-of-Thought Reasoning, Multimodal LLMs, Visual Reasoning, Tool-Calling, Agentic AI, Plug-and-Play Frameworks, Benchmark Evaluation, HPC

- **Indian Institute of Technology (IIT)** — Jodhpur, India
  Researcher - VL2G Labs, Supervisor: Prof. Anand Mishra — *Dec '24 - Aug '25*
  - Developed Model Parity Aligner (MPA) to enhance Small Vision-Language Models (S-VLMs) by systematically identifying and addressing knowledge gaps relative to Large Vision-Language Models (L-VLMs).
  - Designed three key modules (PA, PI, PL) to generate pseudo-annotations, isolate failure cases, and fine-tune S-VLMs on targeted samples, improving efficiency while reducing computational overhead.
  - Achieved notable performance improvements on VQA benchmarks, including +6.4% (TextVQA), +12.0% (ChartQA), and +2.4% (Medical-VQA) on TinyLLaVA-2B. This work was published in **EMNLP Main** in 2025.

  Keywords: Vision-Language Models, Knowledge Distillation, Model Efficiency, Parameter-Efficient Fine-Tuning, Visual Question Answering, Benchmark Evaluation, PyTorch

Honors and Awards

- Awarded the J.N. Tata Endowment Scholarship and Sir Dorabji Tata Trust Travel Grant 2025.
- Graduated with Distinction (both examination and thesis) from IIT Jodhpur, India, ranking in the top 10 of the class.
- Awarded the MITACS GRI 2023 Research Internship.
- Ranked in the top 0.1 percentile in the IIT-JEE Joint Entrance Examination 2020 out of 1.3 million applicants.

Skills Summary

- **Languages**: Python, C++
- **Frameworks & Libraries**: PyTorch, Keras, JAX, HuggingFace Transformers, PEFT (LoRA), TypeScript, SQLite
- **ML & DL Expertise**: Multimodal Learning, Generative Modelling, Vision-Language Models, Audio-Visual Systems, Model Scaling, Knowledge Distillation, Bias Mitigation, Prompt Engineering, LLM Evaluation, Agentic AI, Tool-calling
- **Research Practice**: ML Experimentation, Model Benchmarking, Academic Writing, Peer Review (CVPR, ECCV, NeurIPS)