

ASHIM DAHAL

Undergraduate Researcher in Computer Vision & Multimodal Learning

codeashim@gmail.com | linkedin.com/in/ashimdahal | github.com/ashimdahal

[Google Scholar](#) | ashimdahal.github.io

RESEARCH INTERESTS

Computer Vision, Vision-Language Models, 3D Reconstruction, Video Understanding, Multimodal Learning, Neural Rendering (Gaussian Splatting, NeRFs), Efficient Deep Learning, Vision Transformers

EDUCATION

University of Southern Mississippi

Aug 2023 – May 2027 (Expected)

Hattiesburg, MS, USA

Bachelor of Science in Computer Science; GPA: 3.87/4.0

– Keystone Honors Scholar (merit-based full-tuition scholarship)

– **Relevant Coursework:** Deep Learning, Computer Vision, Machine Learning, Data Structures & Algorithms, Linear Algebra, Probability & Statistics, Software Engineering

RESEARCH EXPERIENCE

Cyber Innovations Lab, University of Southern Mississippi

Aug 2023 – Present

Hattiesburg, MS, USA

Undergraduate Research Assistant, Advisor: Dr. Nick Rahimi

- **Video Question Answering Research:** Developed POVQA (Preference-Optimized Video Question Answering), a novel system that reduces input tokens by 40% while maintaining 95%+ accuracy on long-context video understanding. Implemented rationale generation using vision transformers to enable data-efficient learning on video QA benchmarks.
- **Image Captioning Evaluation Framework:** Created Redemption Score, a multi-modal evaluation metric that triangulates distributional, perceptual, and linguistic signals for robust image-text alignment assessment. Framework outperformed traditional metrics (BLEU, METEOR, CIDEr) on human correlation benchmarks.
- **Vision-Language Model Analysis:** Led comprehensive ablation studies on CLIP and Vision Transformers, analyzing embedding shifts under various augmentation strategies. Research resulted in accepted paper at CVPR 2025 Workshops track.
- **Remote Sensing Segmentation:** Conducted heuristic comparison of Vision Transformers against CNNs for semantic segmentation on satellite imagery (iSAID dataset). Developed novel weighted loss function enabling smaller CNN models to match performance of 5x larger ViT architectures. Published in IEEE Sensors Journal.
- **Distributed Training Optimization:** Architected and optimized 6-GPU PyTorch Distributed Data Parallel training pipeline with automatic mixed precision (AMP) and throughput tuning, achieving 65% reduction in wall-clock time for model training and enabling 3x faster experimental iteration cycles.
- **NASA EPSCoR Project:** Co-designed research methodology for \$51,000 NASA EPSCoR-funded project on efficient vision-language models (PI: Dr. Nick Rahimi). Drafted technical proposal defining objectives, evaluation protocols, and deployment strategies for tiny VLM architectures. Currently leading implementation and experimental evaluation.
- **3D Reconstruction Research:** Awarded \$5,500 DCUR Summer Research Grant (2025) as principal student investigator to develop real-time multiview stereo reconstruction pipeline integrated with dynamic 3D Gaussian Splatting. Implementing custom CUDA kernels for splat rasterization and optimizing differentiable rendering for 30+ FPS novel view synthesis.
- **Neural Architecture Analysis:** Conducted comprehensive study of Convolutional Kolmogorov-Arnold Networks across ImageNet (AlexNet), MNIST (LeNet), and tabular classification tasks. Identified critical efficiency bottlenecks in learnable activation functions, documenting 2.3x training slowdown compared to standard CNNs with minimal accuracy improvements. Published optimization recommendations for future KAN architectures.

Data Research Council for Students

Jun 2022 – Aug 2023

Machine Learning Researcher & Technical Instructor

Kathmandu, Nepal

- **Computer Vision Tool Development:** Built 6 production-ready computer vision applications with RESTful FastAPI endpoints, including GAN-based image enhancement pipeline for NASA JunoCam spacecraft imagery. Project earned Best Local Project with Global Nomination at NASA Space Apps Challenge 2022.
- **Cybersecurity Research:** Developed zero-day cyber attack detection system using multi-layer perceptrons with custom weighted loss functions and SHAP explainability framework. Achieved 94% accuracy on highly imbalanced network intrusion datasets through strategic oversampling and focal loss implementation.

- **Educational Outreach:** Designed and delivered 5 comprehensive Python and Machine Learning bootcamps to 350+ students across Nepal (92% satisfaction rating). Curriculum covered convolutional neural networks, PyTorch distributed training, production deployment with FastAPI, and deep learning fundamentals.
- **Low-Resource NLP:** Developed first Romanized Nepali conversational chatbot using fine-tuned BlenderBot (1.3B parameters) targeting accessibility for low-resource language speakers. Published research in Europe PMC analyzing efficacy of native language interfaces for mental health conversational support.

PUBLICATIONS

Metrics: 9 papers (5 first-author), 24+ citations (Google Scholar), h-index: 3

Peer-Reviewed Conference & Journal Papers

- [1] A. **Dahal**, S. A. Murad, and N. Rahimi, “Embedding shift dissection on clip: Effects of augmentations on vlms’ representation learning,” in *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Proceedings Track, IEEE, 2025.
- [2] A. **Dahal**, S. A. Murad, and N. Rahimi, “Heuristic comparison of vision transformers against convolutional neural networks for semantic segmentation on remote sensing imagery,” *IEEE Sensors Journal*, vol. 25, pp. 17364–17373, 10 2025, **Impact Factor: 4.5**.
- [3] S. A. Murad, A. **Dahal**, and N. Rahimi, “Multilingual cyber threat detection in tweets/x using ml, dl, and llm: A comparative analysis,” *IEEE Transactions on Computational Social Systems*, 2025, **Impact Factor: 4.9**.
- [4] A. **Dahal**, P. Bajgai, and N. Rahimi, “Analysis of zero day attack detection using mlp and xai,” in *International Conference on Security and Management (SAM)*, Las Vegas, NV, USA, 2024.

Preprints & Manuscripts Under Review

- [5] A. **Dahal**, A. Ghimire, S. A. Murad, and N. Rahimi, *Povqa: Preference-optimized video question answering with rationales for data efficiency*, arXiv preprint arXiv:2510.01009, Under Review, 2025.
- [6] A. **Dahal**, A. Ghimire, S. A. Murad, and N. Rahimi, *Redemption score: A multi-modal evaluation framework for image captioning via distributional, perceptual, and linguistic signal triangulation*, arXiv preprint arXiv:2505.16180, 2025.
- [7] A. **Dahal**, S. A. Murad, and N. Rahimi, *Efficiency bottlenecks of convolutional kolmogorov-arnold networks: A comprehensive scrutiny with imagenet, alexnet, lenet and tabular classification*, arXiv preprint arXiv:2501.15757, 2025.
- [8] S. A. Murad, A. **Dahal**, and N. Rahimi, *Eeg-to-text translation: A model for deciphering human brain activity*, arXiv preprint arXiv:2505.13936, 2025.
- [9] A. **Dahal**, A. Khadka, B. Kharal, and A. Shah, *Effectiveness of native language for conversational bots*, Europe PMC Preprints, 2022. doi: PPR561933.

SELECTED TECHNICAL PROJECTS & OPEN SOURCE CONTRIBUTIONS

Torchy – PyTorch Training Utilities Library 2023

- Developed lightweight PyTorch wrapper library providing clean training utilities (logging, checkpointing, mixed precision, metric tracking) while preserving native `nn.Module` workflows and enabling framework-agnostic rapid prototyping.
- Published to PyPI with comprehensive documentation and example implementations. Adopted by open-source community with 15 stars, 5 forks on GitHub.
- **Technologies:** PyTorch, Python, Weights & Biases integration
- <https://github.com/ashimdahal/torchy>

ViT vs. CNN Segmentation Benchmark 2024

- Implemented comprehensive benchmark suite comparing Vision Transformer and CNN architectures for semantic segmentation tasks. Developed novel combined weighted loss function for improved performance on imbalanced remote sensing datasets.

- Structured experiments for reproducibility with detailed ablation studies. Repository directly supports findings published in IEEE Sensors Journal paper (6 stars, 2 forks).
- **Technologies:** PyTorch, Vision Transformers, U-Net, DeepLabV3+
- <https://github.com/ashimdahal/ViT-vs-CNN-Image-Segmentation>

Jelly – Low-Resource Language Chatbot

2022

- Created first Romanized Nepali conversational AI using fine-tuned BlenderBot (1.3B parameters) with Flask web interface, targeting accessibility for 30M+ speakers who use Roman script for digital communication.
- Conducted user study analyzing preference for native vs. English language interfaces in mental health contexts. Published findings in Europe PMC.
- **Technologies:** Hugging Face Transformers, Flask, PyTorch, NLP
- <https://github.com/ashimdahal/jelly>

INVITED TALKS & PRESENTATIONS

“What is an Image?” – Guest Lecture, Introduction to Computer Science

Jan 2025

- University of Southern Mississippi, Hattiesburg, MS
- Introduced foundational concepts of digital images, color spaces, and image processing to 60+ undergraduate students

“Using AI in Journalism” – Professional Development Workshop

2023

- Federation of Nepali Journalists, Kaski, Nepal
- Presented practical applications of AI/ML for news verification, content generation, and investigative journalism

“Future of Artificial Intelligence” – Career Guidance Seminar

2023

- Fishtail Academy, Pokhara, Nepal
- Discussed emerging trends in AI research and career pathways for high school students

TEACHING & MENTORSHIP EXPERIENCE

Machine Learning Bootcamp Series

2022 – 2023

Lead Instructor, Data Research Council for Students

Nepal

- **7-Day Pokhara ML Bootcamp** (2023): Taught 50+ students deep learning fundamentals, CNNs, and PyTorch basics with hands-on projects
- **14-Day Intermediate Python Bootcamp** (2023): Covered advanced Python, object-oriented programming, and ML library usage
- **30-Day Beginner Python Bootcamp** (2022): Introduced 100+ students to programming fundamentals and basic data analysis
- **Chitwan ML Bootcamp** (2023): Delivered intensive week-long training on computer vision and neural networks
- Maintained comprehensive GitHub repositories with course materials, assignments, and project templates for all bootcamps
- Achieved 92% average satisfaction rating across all programs with 350+ total participants

Robotics Instructor

2021 – 2022

Durbar High School

Kathmandu, Nepal

- Designed and taught robotics curriculum covering Arduino programming, sensor integration, and mechanical design
- Mentored student teams in regional robotics competitions

TECHNICAL SKILLS

Programming Languages: Python, C++, C#, SQL, JavaScript, CUDA, Bash

ML/DL Frameworks: PyTorch (DDP, Lightning, FSDP), TensorFlow, JAX, Hugging Face Transformers, Accelerate, scikit-learn, XGBoost

Computer Vision: Vision Transformers (ViT, CLIP, Swin), 3D Gaussian Splatting, Diffusion Models (Stable Diffusion, DALL-E), Neural Radiance Fields (NeRF, Instant-NGP), Video Understanding, Image Segmentation (U-Net, Mask R-CNN), Object Detection (YOLO, Faster R-CNN), Multiview Stereo (MVS)

MLOps & Deployment: Docker, Kubernetes, Git/GitHub, Weights & Biases, MLflow, FastAPI, Flask, Gradio, Streamlit, REST APIs

Cloud & Infrastructure: AWS (EC2, S3, SageMaker), CUDA Programming, Multi-GPU Training, SLURM, High-Performance Computing (HPC)

Libraries & Tools: NumPy, Pandas, OpenCV, Matplotlib, Plotly, Jupyter, Linux/Unix, Neovim, LaTeX, Weights & Biases

Specialized Techniques: Distributed Training (DDP, FSDP), Mixed Precision Training (AMP), Model Optimization, Hyperparameter Tuning, Data Augmentation, Transfer Learning, Few-Shot Learning

HONORS & AWARDS

\$5,500 Summer Research Grant – Drapeau Center for Undergraduate Research	2025
– Competitive merit-based research funding for 3D Gaussian Splatting project (awarded as principal student investigator)	
NASA EPSCoR Research Team Member	2025 – 2026
– Co-designed \$51,000 funded research project on efficient vision-language models	
Keystone Honors Scholar – University of Southern Mississippi	2025 – 2027
– Inducted to the Honors Scholar as a Keystone Scholar for exemplary demonstration of Scholar and Research as an Undergraduate Freshman and Sophomore	
Best Local Project + Global Nomination – NASA Space Apps Challenge	2022
– Developed GAN-based image enhancement tool for JunoCam spacecraft imagery; selected as Nepal's top project	
\$500 Hatchery Checkpoint Award	2024
– Seed funding for XR application development project targeting dyslexia accessibility	
\$200 Eagles Write Award – School of Humanities, USM	2024
– Awarded for best visual analysis essay among all freshman students	

LEADERSHIP & SERVICE

Lead Organizer	2025 – Present
<i>Google Developers Group (GDG) On Campus at University of Southern Mississippi</i>	<i>Hattiesburg, MS</i>
– Organize technical workshops, hackathons, and industry speaker events for student developer community	
– Coordinate with Google Developer Relations team to deliver official GDG curriculum and resources	
Research Liaison	2025 – Present
<i>School of Computing Science & Computer Engineering Student Ambassadors</i>	<i>Hattiesburg, MS</i>
– Represent undergraduate research interests to faculty and administration	
– Connect prospective students with research opportunities and faculty mentors	
Head of Artificial Intelligence	2024 – 2025
<i>Google Developer Students Club (GDSC) at University of Southern Mississippi</i>	<i>Hattiesburg, MS</i>
– Led AI/ML workshop series introducing students to deep learning frameworks and practical applications	
– Mentored 20+ students on machine learning projects and research directions	

PROFESSIONAL AFFILIATIONS

Student Member , IEEE
Member , Google Developer Groups Community

REFERENCES

Available upon request