

Manish Dhakal

Computer Vision . Large Language Models

 [manishdhakal.github.io](https://github.com/manishdhakal)
 mdhakal1@vols.utk.edu
 [Google Scholar](#)
 [manishdhakal](#)

 [manishdhakal521](#)
 [manishdhakal](#)
 Knoxville, TN, USA

Research

Applied deep learning for computer vision and LLMs with the motive of *model compression and efficient learning*. Publications at premier venues, such as *WACV, MICCAI, ACCV, IJCNLP-AACL, CVPR-W* etc.

Education

Ph.D. in Industrial Engineering

University of Tennessee, Knoxville

Graduate School

August 2026 – 2029*

Supervisor: [Yi Ding, Ph.D.](#)

- Research in *efficient machine learning* and *model compression*.

Master's in Computer Science

Georgia State University

Graduate School

August 2024 – July 2026

- Deep learning projects on *2D/3D computer vision and large language models (LLMs)*, funded by the US Department of Defense (DoD).
- Grade: 3.96

Bachelor in Computer Engineering

Pulchowk Campus, Institute of Engineering, Tribhuvan University

Undergraduate School

November 2017 – April 2022

Thesis Supervisor: [Prof. Subarna Shakya](#)

- *Ranked 11th (top 0.1%)* in the entrance exam, competing with *15,000+ candidates*, received a full scholarship for undergraduate studies.
- Gained knowledge about *significant CS courses* like AI, Image Processing, Data Structure & Algorithm, DBMS, Software Engineering, and so on.
- *Thesis*: Automatic speech recognition for *low-resourced Nepali language* which was later presented at an IEEE conference.

Ongoing Projects

Knowledge Distillation

LLMs

- Developing a robust knowledge distillation method utilizing a *teacher-student paradigm* to transfer complex knowledge efficiently.
- Engineering a feature-transfer mechanism between the teacher and student models' intermediate layers to capture and leverage rich, internal representations, ensuring robustness and high fidelity in model compression and knowledge transfer.

Publications

2022-Present

[Scholar](#) Citations: 130+ & h-index: 4

- Poudel, S., Kunwor, N., Simkhada, R., Munir, M., [Dhakal, M.](#), & Poudel, K. (2026, June). Parameter efficient fine-tuning for domain-specific gastrointestinal disease recognition. In *PHAROS @ CVPR*. [[Paper](#), [Code](#)]
- [Dhakal, M.](#), Dasari, V. R., Sunderraman, R., & Ding, Y. (2026, March). GFT: Graph feature tuning for efficient point cloud analysis. In *WACV*. [[Paper](#), [Code](#)]
- [Dhakal, M.](#), Jinadu, U., Budathoki, A., Sunderraman, R., & Ding, Y. (2026, February). DistillLens: Symmetric knowledge distillation through logit lens. *arXiv preprint arXiv:2602.13567*. [[Paper](#), [Code](#)]

- Mim, S. T., Morris, J., Dhakal, M., Xiu, Y., Gorlatova, M., & Ding, Y. (2025, December). Can a unimodal language agent provide preferences to tune a multimodal vision-language model? In *IJCNLP-AACL Findings*. [Paper]
- Budathoki, A., & Dhakal, M. (2025). Adversarial robustness analysis of vision-language models in medical image segmentation. *arXiv preprint arXiv:2505.02971*. [Paper, Code]
- Adhikari, R., Thapaliya, S., Dhakal, M., & Khanal, B. (2024). TuneVLSeg: Prompt tuning benchmark for vision-language segmentation models. In *ACCV (Oral)*. [Paper, Code]
- Dhakal, M., Adhikari, R., Thapaliya, S., & Khanal, B. (2024). Finetuning vision-language segmentation efficiently with lightweight blocks. In *MICCAI*. [Paper, Code]
- Adhikari, R.*, Dhakal, M.*, Thapaliya, S.*, Poudel, K., Bhandari, P., & Khanal, B. (2023, October). Synthetic Boost: Leveraging synthetic data for enhanced vision-language segmentation in echocardiography. In *ASMUS @ MICCAI*. [Paper, Code]
- Poudel, K.*, Dhakal, M.*, Bhandari, P.*, Adhikari, R.*, Thapaliya, S.*, & Khanal, B. (2024). Exploring transfer learning in medical image segmentation using vision-language models. In *MIDL (Oral)*. [Paper, Code]
- Dhakal, M., Chhetri, A., Gupta, A. K., Lamichhane, P., Pandey, S., & Shakya, S. (2022, July). Automatic speech recognition for the Nepali language using CNN, bidirectional LSTM, and ResNet. In *Proceedings of the International Conference on Inventive Computation Technologies (ICICT)*. IEEE. [Paper, Code]

Work Experience

Nepal Applied Mathematics and Informatics Institute for research (NAAMII)

Research Assistant

Lalitpur, Nepal

April 2022 – June 2024

Supervisor: Bishesh Khanal, Ph.D.

- Developed skills for *object detection and segmentation* tasks on 2D medical images and explored their multi-modal approach (esp. *vision-language models*); also worked on segmentation with *3D mesh* data.
- Demonstrated *strong skills in writing scientific manuscripts*, with multiple papers submitted for review, showcasing the ability to *communicate methodologies, results, and implications* effectively.
- Ensured *reproducibility and modularity in ML projects* by implementing robust methodologies and practices, allowing for the transparent and replicable programming of the projects.

Past Projects

Lower Limb Segmentation

Medical Imaging

July 2023 – September 2023

Supervisor: Taman Upadhaya, Ph.D.

- Conducted *training experiments* of different deep learning models on the remote server to segment three bones – knee, pelvis, and ankle – from CT scans of the lower limbs of patients.
- *Deployed a robust Python rest API* on the remote server for the segmentation request from a client, with a pipeline including pre-processing, inference, and post-processing steps.
- Ensured *interoperability, reproducibility, and understandability* of the deployed application using Docker, and well-structured documentation and comments.

Vision Language Segmentation Models (VLSMs) for Medical Images

Medical Imaging

May 2023 – August 2024

- Reported zero-shot and finetuned segmentation performance of *4 VLSMs* on *11 medical datasets* using *9 types of prompts* derived from *14 attributes*, prompts are given as text conditioning information.
- Worked with *encoder-decoder architecture* to generate binary segmentation masks for VLSMs.
- Tested the compatibility of the VLSMs (such as *CLIPSeg and CRIS*) pre-trained for open-domain images with medical images.

Object Detection in 2D Orthopantomogram (OPG) Images

Dental Imaging

September 2022 – August 2024

- Critically analyzed the *literature and state-of-the-art* models for different segmentation and detection tasks on radiology images of dentistry and their inadequacy.
- Designed and developed the *data annotation tool* for object detection over 2D OPG images.
- Working on identification and localization of dental *anatomical structures and abnormalities* while benchmarking with existing methods like *YOLO, RetinaNet, RCNN, and FastRCNN*.

Segmentation in 3D Teeth Scan

Summer 2022

MICCAI Challenge 2022

- Learned about the representation and preprocessing of *3D mesh and point cloud* data.
- Benchmarked with different 3D point cloud segmentation models such as *Pointnet++ and DeltaConv*.

Nepali AutoComplete and LM

August 2020 – October 2020

[Open Source Project](#)

- Designed and trained *language model of Nepali (ie. Devnagari transcript)* for the text auto-complete system.
- Programmed the *pre-processing pipeline* to remove the non-Nepali characters from the dataset.

Super-Resolution with GAN (SRGAN)

May 2020 – August 2020

[Open Source Project](#)

- Implemented *open source model* of *SRGAN* with Keras/TensorFlow.
- Developed the *understanding of generator and discriminator* in GAN-based generative models.

Teaching

Community Eye, ENT & Rehabilitation Center (CEERS)

Trainer

Bhaktapur, Nepal

June 2023 – June 2024

- *Training a group of interns* to develop medical imaging applications with the use of ML.
- Instructing and guiding them about ML through activities like *paper reading sessions, lecture-lab sessions, and topic presentations*.

4th Annual Nepal AI School (ANAIS)

Teaching Assistant

Kathmandu, Nepal

May 2023 – June 2023

- *Guided participants* through a series of labs related to *neural networks, transformers, federated learning, graph neural networks, active learning*, and so on.
- *Mentored three groups* during the 10-day *machine learning hackathon* (namely, Hack-a-Dev).

Software Fellowship, Locus 2021

Programming Instructor

Online

Summer 2021

- Provided tutoring on *software development life cycle* and assisted participants with *software documentation* and *library/framework installation*.
- Taught participants about *API development for web applications*, emphasizing its concepts, best practices, and usage.

Achievements/Services

2026	<i>MICCAI Reviewer</i>
2025	<i>AAAI Reviewer</i>
2024	<i>Presidential Fellowship</i> , GSU (Only 6% of the doctoral students @ GSU).
2024	<i>LMIC Travel Grant</i> by the conference of MICCAI.
2017	<i>Undergraduate</i> funded by the Government of Nepal.

Technical skills

Deep Learning	<i>Unimodal and multimodal learning</i> , large language models (LLMs), model quantization, efficient learning, reinforcement learning fine-tuning, 2D/3D vision, convolution, and transformers.
Writing	<i>Knowledge synthesis</i> from the existing literature, <i>writing scientific documents</i> and <i>manuscripts</i> with <i>LaTeX</i> , and <i>communicating the results</i> to the community with transparency.
Remote Server	Able to work with <i>remote Linux machines</i> for coding and project deployment using <i>SSH, shell script, tmux, Nginx, and Docker</i> .

References (Research Advisors)

Yi Ding, Ph.D.

Assistant Professor, *Department of Industrial and Systems Engineering, University of Tennessee Knoxville*
yding@utk.edu

Raj Sunderraman, Ph.D.

Professor, *Department of Computer Science, Georgia State University (GSU)*
rsunderraman@gsu.edu

Bishesh Khanal, Ph.D.

Research Director, *Nepal Applied Mathematics and Informatics Institute for research (NAAMII)*
bishesh.khanal@naamii.org.np