Manish Dhakal

Computer Engineer . Machine Learning Researcher

+977 986-0687860 | manishdhakal.com.np
manishdhakal521 | manish.dhakal@naamii.org.np
manishdhakal | Google Scholar
Nepal | manishdhakal

Summary

An accomplished computer engineer with experience as a Research Assistant (RA), interested in multimodal learning, medical imaging, and continual learning, skilled in maths and programming, and adept at communicating research results to the community.

Work Experience

Nepal Applied Mathematics and Informatics Institute for research (NAAMII) — *Lalitpur, Nepal*
Research Assistant — *April 2022 – Present*

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 for 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.

Education

Bachelor in Computer Engineering — *Bachelors*
Pulchowk Campus, Institute of Engineering, Tribhuvan University — *November 2017 – April 2022*

Thesis Supervisor: Prof. Subarna Shakya

- *Ranked 11th (top 0.1%)* in the engineering entrance exam, competing with *15,000+ candidates* , received *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.

Publications

Vision-Language Model, NLP, Medical Imaging
Lead author / Presenter / Co-first author — *2022-Present*

- *Dhakal, M.* , Adhikari, R., Thapaliya, S., & Khanal, B. (2024). VLSM-Adapter: Finetuning Vision-Language Segmentation Efficiently with Lightweight Blocks. arXiv preprint arXiv:2405.06196.[1]
- 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 International Workshop on Advances in Simplifying Medical Ultrasound* (pp. 89-99). Cham: Springer Nature Switzerland.[2]
- Poudel, K.*, *Dhakal, M.** , Bhandari, P.*, Adhikari, R.*, Thapaliya, S.*, & Khanal, B. (2023). Exploring Transfer Learning in Medical Image Segmentation using Vision-Language Models. *arXiv preprint* arXiv:2308.07706.[3]
- *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 2022 International Conference on Inventive Computation Technologies (ICICT)* (pp. 515-521). IEEE.[4]

[1] https://github.com/naamiinepal/vlsm-adapter
[2] https://github.com/naamiinepal/synthetic-boost
[3] https://github.com/naamiinepal/medvlsm
[4] https://github.com/manishdhakal/ASR-Nepali-using-CNN-BiLSTM-ResNet

Teaching

Community Eye, ENT & Rehabilitation Center (CEERS) *Bhaktapur, Nepal*
Trainer *June 2023 – Present*

- *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) *Kathmandu, Nepal*
Teaching Assistant *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 *Online*
Programming Instructor *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.

Projects

Lower Limb Segmentation *July 2023 – September 2023*
Medical Imaging

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 *May 2023 – Present*
Medical Imaging

- 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 *September 2022 – Present*
Dental Imaging

- 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.

Technical skills

Machine Learning *Unimodal and multimodal (esp. vision-language model)* ML project structuring for detection and segmentation task while maintaining *reproducibility and modularity* ; integrating *open source models* for benchmarking. Proficiency in using libraries and frameworks like *NumPy, Pandas, PyTorch, and TensorFlow* .

Writing *Knowledge synthesis* from the existing literature, *writing scientific documents* and *manuscripts* with *LaTex* , and *communicating the results* to the community with transparency.

Web Development Competence in creating well-documented backend applications with relational databases using frameworks like *Django, FastAPI, and NodeJS* . Adept at client-side programming with *ReactJS.*

Remote Server Able to work with *remote Linux machines* for coding and project deployment using *SSH, shell script, tmux, Nginx, and Docker* .

Extracurricular Activities

DataRush (AI and Data Science Competition, Locus) *Spring 2022*
Co-ordinator

- *Call for sponsors* , maintained communication and coordination between sponsors, participants, mentors, and other organizers.
- Made the *budget planning* , prepared the *event's rule book* , planned the *event structure* , and ensure the smooth operations of the event.
- *Tested and validated* the machine learning models and their *solutions* submitted by the participants.

Achievements and Awards

Full Scholarship B.E. funded by Government of Nepal.
Winner LogPoint CTF, Cybersecurity Competition.
Finalist Spirathon 2020, TechXperience Nepal.

Certifications

Stanford University Machine Learning
DeepLearning.AI Deep Learning Specialization
DeepLearning.AI AI for Medicine
DeepLearning.AI Natural Language Processing
DeepLearning.AI TensorFlow in Practice
University of Alberta Reinforcement Learning

References

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

Prof. Subarna Shakya
Professor of Computer Engineering, *Department of Electronics and Computer Engineering, Pulchowk Campus, Institute of Engineering, Tribhuvan University*
drss@ioe.edu.np

Taman Upadhaya, Ph.D.
Associate Researcher, *University of California San Francisco* **||** Adjunct Research Scientist, *Nepal Applied Mathematics and Informatics Institute for research (NAAMII)*
taman.upadhaya@naamii.org.np