Rajan Gyawali

Columbia, MO 65201 | rgkg2@missouri.edu | 573-639-7487

linkedin.com/in/rajangyawali | github.com/rajangyawali | scholar.google.com/citations?user=qzXOzEkAAAAJ&hl=en

Education

University of Missouri **Aug 2022 – Dec 2026 (Expected)**
PhD in Computer Science *Columbia, MO*
Emphasis: Deep Learning and Bioinformatics
Advisor: Dr. Jianlin Cheng

Tribhuvan University **2017 – 2019**
Masters of Science in Information and Communication Engineering *Kathmandu, Nepal*
Thesis Advisor: Dr. Dibakar Raj Pant
Thesis: Employee Face Recognition by Region Proposal Networks and Faster R-CNN

Tribhuvan University **2012 – 2016**
Bachelors of Engineering in Electronics and Communication *Kathmandu, Nepal*

Research Experience

University of Missouri **August 2022 – Present**
Graduate Researcher at Bioinformatics and Machine Learning Lab *Columbia, MO*

- Designed and developed MICA. (MICA Code, MICA Paper).
- MICA is a fully automated, multimodal deep learning integration of cryo-EM density maps and AlphaFold3 structures that generates 3D atomic structures of the protein complexes.
- MICA utilizes an encoder-decoder architecture with feature pyramid networks to identify backbone atoms, carbon alpha atoms & amino acid types within 3D cryo-EM density maps with integration of AlphaFold3 structures and then utilizes a backbone tracing protocol within the predicted atoms and amino acid types to form the 3D atomic structures of proteins.
- Designed and developed CryoSegNet. (CryoSegNet Code, CryoSegNet Paper).
- CryoSegNet integrates SAM with attention-gated U-Net that accurately predicts protein particles from noisy cryo-EM micrographs, achieving 7-14% improvement in 3D reconstruction resolution (3.32Å) over established baselines.
- Designed and developed CryoPPP Dataset. (CryoPPP Code, CryoPPP Paper).
- CryoPPP is a curated dataset of preprocessed and labeled cryo-EM micrographs, designed to train CryoSegNet, CryoTransformer, and similar AI models for automated protein particles picking from cryo-EM micrographs. This 2.6TB benchmark dataset (9,893 micrographs, 300K+ labels, 34 proteins) is now a widely adopted standard training/evaluation dataset in community which has already been used in 45+ publications.
- The GitHub repository for CryoPPP contains source code to preprocess and label cryo-EM micrographs for deep learning model training and evaluation.
- Contributed to CryoTransformer (detection transformer with 0.747 F1-score) and CryoFSL (few-shot learning reducing labeling by 95%) for automated protein particles detection in cryo-EM micrographs.
- Trained and experimented with deep learning models (including Graph Neural Networks) using distributed deep learning training approach (Distributed Data Parallel) in multi-node multi-GPUs.
- Developed process-based parallelism scripts to preprocess data in parallel, utilizing all physical cores across multiple nodes in Slurm and LSF clusters.

Technical Skills

Deep Learning Architectures: CNNs, Vision Transformers, U-Net, ResNet, attention mechanisms, encoder-decoder models, Graph Neural Networks (GNNs), E(n)-equivariant networks

ML Frameworks & Libraries: PyTorch, PyTorch Lightning, PyTorch Geometric, Deep Graph Library, scikit-learn

Distributed Computing: Multi-node multi-GPU training (DDP), HPC clusters (Slurm, LSF), parallel data preprocessing

Programming: Python, C++, Bash, R, SQL

Structural Biology Tools: PyRosetta, UCSF ChimeraX, PyMOL, RDKit, Open Babel, Clustal Omega, cryo-EM software (RELION, CryoSPARC)

Development: Git/GitHub, Linux, Docker, Visual Studio, RStudio

Publication

[8] **Gyawali, R.**, Dhakal, A., & Cheng, J. (2025). **Multimodal deep learning integration of cryo-EM and AlphaFold3 for high-accuracy protein structure determination.** *Communications Chemistry, 8(1), 320.*

[7] Dhakal, A.*, **Gyawali, R.***, Wang, L., & Cheng, J. (2025). **Artificial intelligence in cryo-EM protein particle picking: recent advances and remaining challenges.** *Briefings in Bioinformatics, 26(1).*

[6] He, F., Yang, Z., Gao, M., Poudel, B., Dhas, N.S.E.S., **Gyawali, R.**, Dhakal, A., et al. (2024). **Adapting Segment Anything Model (SAM) through Prompt-based Learning for Enhanced Protein Identification in Cryo-EM Micrographs.** *2024 IEEE International Conference on Medical Artificial Intelligence (MedAI).*

[5] **Gyawali, R.**, Dhakal, A., Wang, L., & Cheng, J. (2024). **CryoSegNet: accurate cryo-EM protein particle picking by integrating the foundational AI image segmentation model and attention-gated U-Net.** *Briefings in Bioinformatics, 25(4).*

[4] Dhakal, A., **Gyawali, R.**, Wang, L., & Cheng, J. (2024). **CryoTransformer: a transformer model for picking protein particles from Cryo-EM micrographs.** *Bioinformatics, 40(3), btae109.*

[3] **Gyawali, R.***, Dhakal, A., Wang, L., & Cheng, J. (2023). **CryoVirusDB: a labeled cryo-EM image dataset for AI-driven virus particle picking.** *bioRxiv, 2023.12.25.573312.*

[2] Dhakal, A.*, **Gyawali, R.***, Wang, L., & Cheng, J. (2023). **A large expert-curated cryo-EM image dataset for machine learning protein particle picking.** *Scientific Data, 10(1), 392.*

[1] **Gyawali, R.**, & Pant, D.R. (2019). **An Approach for the Employee Face Recognition by RPN and Faster R-CNN Techniques**. *Proceedings of IOE Graduate Conference, 2019-Summer, 6 (ISSN 2350-8914), 231.*

**Equal contribution*

Other Experiences

Nepal Telecom | Government of Nepal **May 2018 – June 2022**
Telecommunications Engineer *Kathmandu, Nepal*

- Developed "Wireless KPI Analytics Tool" - a fully automated Django-based analytics system for analyzing key performance indicators of network elements eliminating manual analysis.
- Scripted automation for daily redundant tasks related to report management and network status.
- Prepared estimates, installation plans, performance reports, and documentation for different projects.
- Provided training to junior employees on "Python Programming" and "Data Analytics and Visualizations."

Himalaya College of Engineering, Tribhuvan University **November 2016 – May 2018**
Lecturer *Kathmandu, Nepal*

- Instructed undergraduate courses on Artificial Intelligence, Computer Graphics, C Programming, Python Programming and Instrumentation Systems.
- Prepared course materials including lab experiments, lectures, exams, tutorials and assignments.

Talks & Presentations

CryoSegNet: AI-based method for proteins particle picking from cryo-EM density micrographs **2024**
Cryo-EM Super Group *University of Missouri*

A large expert-curated cryo-EM image dataset for machine learning protein particle picking **2024**
RECOMB Conference (poster presentation) *Cambridge, MA*

Honors and Awards

Dean's Fellowship, University of Missouri 2022

Graduate Scholarship, Tribhuvan University 2017

Undergraduate Scholarship, Tribhuvan University 2012

Professional Service and Other Involvements

Manuscript Reviewer:

NeurIPS, Briefings in Bioinformatics, BMC Bioinformatics, Communications Chemistry

President, University of Missouri Nepalese Students Association May 2025 – April 2026

Treasurer, University of Missouri Nepalese Students Association May 2024 – April 2025

Member, Academic Affairs Committee, University of Missouri August 2023 – May 2024