Subash Khanal

subash.khanal.cs@gmail.com Personal Website Google Scholar

EDUCATION

Washington University in St. Louis — **St. Louis, MO, U.S.**
Doctor of Philosophy in Computer Science — *2022 - 2025*

University of Kentucky — **Lexington, KY, U.S.**
Master of Science in Electrical Engineering (GPA: 4.0/4.0) — 2018 - 2020

Nitte Meenakhshi Institute of Technology — **Bengaluru, India**
Bachelors in Electronics and Communication Engineering (GPA: 9.54/10 - Gold Medalist) — *2012 - 2016*

WORKING EXPERIENCE

Multimodal Vision Research Lab — **Washington University in St. Louis**
Postdoctoral Research Associate — *August 2025 - Present*

- Develop multi-scale, multi-condition satellite image generative models for simulating Earth's surface.
- Develop deep learning models for agricultural field delineation from satellite imagery.
- Design multimodal representation learning frameworks for diverse Geospatial AI tasks.
- Mentor graduate students on multimodal learning research projects.

Multimodal Vision Research Lab — **Washington University in St. Louis**
Graduate Research Assistant — *August 2022 - August 2025*

- Developed multi-modal deep learning frameworks for geospatial understanding of global soundscapes.
  - **Sat2Sound** (Under Review 2025): Developed a state-of-the-art soundscape mapping framework that leverages a Vision-Language Model (VLM) to enrich the semantic understanding of a location's soundscape, learns a shared codebook for fine-grained alignment, and enables retrieval-based, location-conditioned soundscape generation.
  - **PSM** (ACM Multimedia 2024): Extended GeoCLAP with a probabilistic, multi-scale, and metadata-aware embedding space for improved zero-shot soundscape mapping.
  - **GeoCLAP** (BMVC 2023): Proposed a tri-modal embedding space integrating satellite imagery, audio, and textual descriptions for zero-shot soundscape mapping at a global scale.
- Built large-scale ML-ready datasets:
  - **GeoSound**: A multi-resolution dataset combining over 300k paired samples of satellite imagery, geotagged sounds, and textual sound descriptions.
  - **EarthCaps**: A global dataset comprising ∼2.5 million satellite images and ∼10 million auditory and visual captions, covering the entire Earth, generated using a state-of-the-art multimodal LLM.
- Contributed to **RANGE** (CVPR 2025), a novel retrieval-augmented method for multi-resolution geolocation embeddings.
- Contributed to additional multimodal embedding spaces:
  - **Sat2Cap** (EarthVision 2024) and **GeoBind** (IGARSS 2024) for query-driven mapping.
  - **TaxaBind** (WACV 2025) for ecological applications.
- Contributed to frameworks for ecological modeling and visual classification:
  - **LD-SDM** and **BirdSAT** (WACV 2024) for species distribution modeling and fine-grained classification.
- Contributed to diffusion model-based conditional generation methods:
  - **GeoSynth** (EarthVision 2024): A framework for controlled satellite image synthesis.
  - **MVPS** (Accepted to TMLR 2025): A framework for novel panoramic view synthesis using satellite imagery.
- Collaborated on the IARPA-funded **SMART** project:
  - Developed large-scale self-supervised learning (SSL) methods, including Masked Autoencoders (MAE), for remote sensing applications such as semantic change detection.

Valuation and Market Dynamics — **Zillow Group**
Applied Scientist - Intern — *May 2024 - August 2024*

- Explored directions for improving Zillow's flagship product, Zestimate, by incorporating a diverse set of features.
- Developed a multimodal model for sales price estimation, learning from both structured tabular data and floorplan images.
- The proposed multimodal model achieved approximately 3% improvement across all metrics compared to the baseline.

Lin Brain Lab — **University of Kentucky**
Graduate Research Assistant — *August 2020 - August 2022*

- Provided applied Machine Learning (ML) and data science support to advance Alzheimer's disease (AD) research while working on different modalities such as medical imaging, electronic health records, and genomics.
- Designed CNN and Vision Transformers (ViT) based models trained on MRI/PET imagery for AD prediction.
- Focusing on interpretability, designed CAT-XPLAIN, an inherently interpretable ViT model that uses causality in predictions to identify the most important image patches.
- Built ML models trained on genetics, electronic health records, and imaging features for biomarker discovery and early prediction of Alzheimer's disease.

Speech and Signal Processing Lab **University of Kentucky**
Graduate Research Assistant *August 2018 - August 2020*

- Performed analysis of articulatory differences in the speech of native and non-native speakers of English.
- Built an Automatic Speech Recognition (ASR) based Mispronunciation Detection and Diagnosis (MDD) framework. ASR was trained using Recurrent Neural Networks (RNNs) on both articulatory and acoustic features.

Kantipur Engineering College **Lalitpur, Nepal**
Lecturer *April 2017 - July 2018*

- Taught courses: Microprocessor, Instrumentation
- Designed and conducted a lab on Digital Signal Processing (DSP), Microprocessor.

KEY SKILLS

- **Languages and Tools:** Python (fluent), Pytorch (fluent), Git (fluent), Docker (familiar), QGIS (familiar).
- **Research Interests and Relevant Courses:** Computer Vision, Multimodal Machine Learning, Deep Learning, Self-Supervised Learning, Generative AI, Geospatial AI, Data Science, Data Structures and Algorithms.

AWARDS AND ACHIEVEMENTS

- COMPEX Scholarship offered by the Indian Embassy in Nepal for undergraduate study in India, 2012–2016.
- ECE Gold Medal, 2016 for graduating with Rank 1 in the department.

PUBLICATIONS

Preprints:

- **Khanal Subash**, Sastry Srikumar, Dhakal Aayush, Ahmad Adeel, and Jacobs Nathan, "Sat2Sound: A Unified Framework for Zero-Shot Soundscape Mapping," arXiv:2505.13777, 2025.
- Sastry Srikumar, Xin Xing, Dhakal Aayush, **Khanal Subash**, Ahmad Adeel, and Jacobs Nathan, "LD-SDM: Language-Driven Hierarchical Species Distribution Modeling," arXiv:2312.08334, 2024.
- **Khanal Subash**, Brodie Benjamin, Xing Xin, Lin Ai-Ling, and Jacobs Nathan, "Causality for inherently explainable transformers: CAT-XPLAIN," **Spotlight Presentation** at XAI4CV: Explainable Artificial Intelligence for Computer Vision (IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops), arXiv:2206.14841, 2022.

Published Papers:

- Sastry Srikumar, Dhakal Aayush, Xing Eric, **Khanal Subash** and Jacobs Nathan, "Global and Local Entailment Learning for Natural World Imagery, " International Conference on Computer Vision (ICCV), 2025.
- Xiong Zhexiao, Xing Xin, Workman Scott, **Khanal Subash** and Jacobs Nathan, "Mixed-View Panorama Synthesis using Geospatially Guided Diffusion, " Transactions on Machine Learning Research(TMLR), 2025.
- Dhakal Aayush, Sastry Srikumar, **Khanal Subash**, Ahmad Adeel, Xing Eric, and Jacobs Nathan, "RANGE: Retrieval Augmented Neural Fields for Multi-Resolution Geo-Embeddings, " IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2025.
- Sastry Srikumar, **Khanal Subash**, Dhakal Aayush, Ahmad Adeel, and Jacobs Nathan, "TaxaBind: A Unified Embedding Space for Ecological Applications," IEEE Winter Conference on Applications of Computer Vision (WACV), 2025.
- **Khanal Subash**, Xing Eric, Sastry Srikumar, Dhakal Aayush, Xiong Zhexiao, Ahmad Adeel, and Jacobs Nathan, "PSM: Learning Probabilistic Embeddings for Multi-scale Zero-Shot Soundscape Mapping," ACM Multimedia, 2024, doi: 10.1145/3664647.3681620.
- Dhakal Aayush, Ahmad Adeel, **Khanal Subash**, Sastry Srikumar, Kerner Hanna,h and Jacobs Nathan, "Sat2Cap:

Mapping Fine-Grained Textual Descriptions from Satellite Images," IEEE/ISPRS Workshop: Large Scale Computer Vision for Remote Sensing (EARTHVISION), 2024. **Best Paper Award**.

- Sastry Srikumar, **Khanal Subash**, Dhakal Aayush, and Jacobs Nathan, "GeoSynth: Contextually-Aware High-Resolution Satellite Image Synthesis," IEEE/ISPRS Workshop: Large Scale Computer Vision for Remote Sensing (EARTHVISION), 2024.
- Dhakal Aayush, **Khanal Subash**, Sastry Srikumar, Ahmad Adeel, and Jacobs Nathan, "GeoBind: Binding text, image, and audio through satellite images," IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 2024. **Oral Presentation**.
- Sastry Srikumar, **Khanal Subash**, Dhakal Aayush, Di Huang, and Jacobs Nathan, "BirdSAT: Cross-View Contrastive Masked Autoencoders for Bird Species Classification and Mapping," IEEE Winter Conference on Applications of Computer Vision (WACV), 2024.
- **Khanal Subash**, Sastry Srikumar, Dhakal Aayush, and Jacobs Nathan, "Learning Tri-modal Embeddings for Zero-Shot Soundscape Mapping," British Machine Vision Conference (BMVC), 2023.
- Xing Xin, Liang Gongbo, Zhang Yu, **Khanal Subash**, Lin Ai-Ling, and Jacobs Nathan, "Advit: Vision transformer on multi-modality pet images for alzheimer disease diagnosis," IEEE International Symposium on Biomedical Imaging (ISBI), 2022, doi: 10.1109/ISBI52829.2022.9761584.
- **Khanal Subash**, Chen Jin, Jacobs Nathan, and Lin Ai-Ling, "Alzheimer's Disease Classification Using Genetic Data," IEEE International Conference on Bioinformatics and Biomedicine (BIBM) Workshop, 2021, doi: 10.1109/BIBM52615.2021.9669730.
- Brodie Benjamin, **Khanal Subash**, Rafique Muhammad Usman, Greenwell Connor, and Jacobs Nathan, "Hierarchical Probabilistic Embeddings for Multi-View Image Classification," IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 2021, doi: 10.1109/IGARSS47720.2021.9554405.
- **Khanal Subash**, Johnson Michael T., Soleymanpour Mohammad, and Bozorg Narjes, "Mispronunciation Detection and Diagnosis for Mandarin Accented English Speech," International Conference on Speech Technology and Human-Computer Dialogue (SpeD), 2021, doi: 10.1109/SpeD53181.2021.9587408.
- **Khanal Subash**, Johnson Michael T. and Bozorg Narjes. "Articulatory Comparison of L1 and L2 Speech for Mispronunciation Diagnosis," IEEE Spoken Language Technology Workshop (SLT), 2021, doi: 10.1109/SLT48900.2021.9383574.
- **Khanal Subash**, "Mispronunciation Detection and Diagnosis for Mandarin Accented English Speech," Theses and Dissertations–Electrical and Computer Engineering, 156, 2020, doi: 10.13023/etd.2020.340.