BISHWASH KHANAL

khanal.bishwash08@gmail.com ◇ bkhanal-11 ◇ Linkedin ◇ ORCID ◇ Jyväskylä, Finland

EDUCATION

Master's Degree Programme (M.Sc.) in Artificial Intelligence 08.2024 - present
University of Jyväskylä Jyväskylä, Finland
Relevant Courses: Deep Learning for Cognitive Computing, Semantic Web and Linked Data, SOA and Cloud Computing, Collective Intelligence and Agent Technology, Computer Vision and Image Analysis, Natural Language Processing, and Simulation.

Bachelor of Science in Electrical and Computer Engineering 09.2017 - 08.2020
Jacobs University Bremen Bremen, Germany
Minor in **Intelligent Mobile Systems**; Grade: 1.73 (German grading system)
Relevant Courses: Signal Processing, Wireless Communication, Information Theory, Machine Learning, Computer Vision, Artificial Intelligence, Robotics, Embedded Systems.
Thesis: *Non-Linearity in Wireless Communications and Deep Learning*.

PROFESSIONAL EXPERIENCE

Artificial Intelligence Engineer Mar 2024 - present
OptiML Org (part-time, remote) California, USA

- **LLM Finetuning and Compression** 2024-present
  This project involves integrating powerful model compression techniques into fine-tuning workflows, empowering AI developers and researchers to optimize and deploy large models with efficiency and scale. Founding member of the team.
  - **My Responsibilities:** Develop a custom pipeline for fine-tuning and compressing open-source LLM models together, evaluating their performance on domain-specific tasks; research and evaluation of SOTA LLM Compression techniques.
  - **Tools Used:** Python (torch, transformers, datasets, peft, bitsandbytes, accelerate, wandb)
  - **Models Used:** Llama, Qwen, Mistral, Gemma. Phi, Deepseek

Computer Vision Engineer Mar 2021 - Aug 2024
E.K. Solutions Pvt. Ltd. (full-time, on-site) Lalitpur, Nepal

- **Model Compression** 2024
  This project involved the quantization of computer vision models for face detection, recognition, and swapping to enable real-time performance on edge devices like the Jetson AGX Orin.
  - **My Responsibilities:** Set up and deploy face detection, recognition, and swapping models on Jetson AGX Orin; perform `int8` model quantization using PyTorch Quantization and NVIDIA TensorRT for optimized and real-time performance.
  - **Tools Used:** Python (numpy, opencv, ultralytics, tensorrt, onnx, insightface)
  - **Models Used:** YOLOv8, Insightface (ARCFace, Inswapper), Hyperstyle (StyleGAN2, CLIP), SOLIDER
- **LLM Tools** 2023-2024
  The project involved the development of a suite of multifaceted tools integrating Large Language Models (LLMs) for enhanced software project management, coding, documentation, meeting and video analysis, automation, QA, HR, and administrative tasks.
  - **My Responsibilities:** Development of a chatbot with a workspace containing user stories, transcriptions, and database analysis using Retrieval-Augmented Generation (RAG); integration of bulk Curriculum Vitae (CV) analysis.
  - **Tools Used:** Python (langchain, openai, sqlalchemy, chromadb, fastapi, pandas)
  - **Models Used:** GPT-4 (through API), Sentence Transformers
- **Virtual Tours** 2022-2024
  The project was about the generation of planar simplified 3D mesh for virtual tours of multi-room indoor scenes captured with iPad Pro Lidar, followed by texture projection using MVS texturing.
  - **My Responsibilities:** Geometrical analysis, formulation, and implementation of algorithms for planar simplification and texturing; implementation of ICP registration and TSDF volumetric integration for combining multiple 3D meshes; application of hole filling, mesh simplification, and texturing enhancement for 3D meshes.
  - **Tools Used:** Python (numpy, opencv, open3d, fastapi, celery), Meshlab, C++, Swift5, Javascript (three.js)
  - **Models Used:** RandLA-Net (Semantic Segmentation)
- **Landmark Detection** 2021
  The goal of the project was to develop a custom object detection model that detects faulty knot-bolt arrangements on large construction sites.
  - **My Responsibilities:** Labeling data for training deep Convolutional Neural Networks and YOLO models; fine-tuning pre-trained YOLO models based on new data and requirements.
  - **Tools Used:** Python (numpy, tensorflow)

  - **Models Used:** YOLOv5, CNN

- **Facial Check-In System** 2021
  This project involved the development of a system that would handle daily employee check-in/check-out through face detection.
  - **My Responsibilities:** Fine-tuning FaceNet models with an improved loss function; developing a pipeline for matching facial embeddings with the existing database using Naive Bayes Algorithm.
  - **Tools Used:** Python (numpy, pytorch, scipy, scikit-learn)
  - **Models Used:** FaceNet

Competitive Coder and Data Analyst 06.2023 - 10.2023
Scale AI (freelancing, remote) California, USA
Worked as a Freelancer through Upwork that involved data annotations to help train Generative AI models to become better writers.

- **My Responsibilities:** Solving competitive programming problems to train AI model; ranking a series of responses that were produced by an AI model; writing a short story based on a provided topic; assessing whether a piece of text produced by an AI model is factually accurate or not.

Computer Vision Engineer 04.2023
GoThru Media Inc. (part-time, remote) Ontario, Canada
Worked as a part-time Computer Vision Engineer to help automate pose estimation between spherical panoramas which was otherwise done manually.

- **My Responsibilities:** Estimating pose between uncalibrated unordered set of spherical panoramas.

Kitchen Staff 02.2019 - 02.2021
Hegartys Irish Pub (part-time, on-site) Bremen, Germany
Worked as a part-time kitchen assistant. My responsibilities included basic food preparation, meal presentation, stocking kitchen and storeroom, and cleaning duties.

ACADEMIC EXPERIENCE

Student Supervisor 07.2018 - 08.2020
Jacobs University Bremen Bremen, Germany
Worked at the Information Resource Center (IRC)/Library as a Student Assistant. Tasks included supporting fellow students, guests, and patrons around the library; training and assisting new employees; checking in and out books and multimedia items; processing newspapers and journals; shelving books; and opening and closing the library.

Teaching Assistant, Intro to Electronics 02.2020 - 05.2020
Jacobs University Bremen Bremen, Germany
Tasks included helping students by providing guidance on weekly assignments and mentoring in software tools like KiCAD and LTSpice for PCB simulations; assessing exams and assignments, ensuring accurate grading, and conducting tutorial sessions to further support student comprehension.

Teaching Assistant, Communications Lab 09.2019 - 12.2019
Jacobs University Bremen Bremen, Germany
Tasks included aiding students in programming with C and MATLAB to simulate Digital Signal Processing and Communications systems; evaluating and grading lab reports and exams; and conducting tutorials to reinforce understanding and facilitate learning.

SKILLS

Programming Languages	Python, C, C++, MATLAB, Java	**ML Tools**	Tensorflow, Keras, PyTorch, Scikit-Learn
ML Models	LLMs, Transformers, GANs, CNNs	**CI/CD Tools**	Docker, Kubernetes
API Frameworks	FlaskAPI, FastAPI, Jakarta	**Database**	MySQL, Postgre SQL
Environments	Linux, Unix, Windows	**Miscellaneous Tools**	Git, GitHub, LaTeX, VS Code

PUBLICATIONS

- B. Khanal, M. Om, S. Rijal, V. Ojha (2024), Alignment of a 360° image with posed color images for locally accurate texturing of 3D mesh, *Frontiers in Computer Science*, Volume 6 - 2024, ISSN: 2624-9898.
- B. Khanal, S. Rijal, M. Awale, V. Ojha (2024), Structure-preserving Planar Simplification for Indoor Environments, *arXiv*.
- B. Khanal and J. M. Capone, (2024), Evaluating the Impact of Compression Techniques on Task-Specific Performance of Large Language Models, *arXiv*.

SCHOLARSHIPS

- **JYU Scholarship**, University of Jyväskylä 2024-2026
  This scholarship is a university-level scholarship awarded by the JYU rector's decision awarded to students who demonstrate exceptional academic achievement. This scholarship covers 50% of the student's tuition fees.
- **Merit-Based Scholarship**, Jacobs University Bremen 2017-2020
  This scholarship is awarded to students who demonstrate exceptional academic achievement, leadership potential, and a strong

commitment to their field of study. This scholarship recognizes individuals who exhibit outstanding merit and promise in their academic pursuits.

- **Mahatma Gandhi Scholarship**, Embassy of India in Kathmandu 2015-2017
  This scholarship is typically awarded to outstanding students who demonstrate exceptional academic achievements during High School and have a poor economic background.

CERTIFICATIONS

Deep Learning Specialization, Coursera
GAN Specialization, Coursera
MLOps Specialization, Coursera
Meta Database Engineer Specialization, Coursera
NLP Specialization, Coursera
ML Specialization, Coursera
DeepLearning.AI TensorFlow Developer, Coursera

MISCELLANEOUS

- **Nepali OCR using Vision based Transformers** 09.2023
  Inspired by the popular TR-OCR, exploring the possibilities of creating vision-based transformers for detecting Nepali text. The previous works have already been replicated and implemented on the datasets. A RoBERTa based tokenizer and encoder has been trained on Nepali datasets. Exploring on methodologies for collecting datasets for OCR.
  Project Resources: Nepali RoBERTa Model.
- **Face Generator using Generative Models** 03.2023-06.2023
  Explored and applied encoder-decoder frameworks, including autoencoders, VAE, DCGAN, and WGAN, to generate synthetic facial images with a focus on enhancing realism and diversity, utilizing the CelebA dataset. Conducted comprehensive mathematical and visual assessments across various model variants.
  Project Resources: Infusing the Diffusion Models, Applications of GAN.
- **Google - American Sign Language Fingerspelling Recognition** 05.2023-08.2023
  Participated in a competition organized by Google and hosted in Kaggle. Task included detection and translation American Sign Language (ASL) fingerspelling into text. Worked on a TensorFlow Lite model trained on labeled landmark data extracted using the MediaPipe Holistic Solution. The model achieved private score of 0.647 and public score of 0.696 which was able to secure 371th ranking out of 1315 teams.
  Project Resources: Google - ASL Fingerspelling Recognition, Google - ISL Recognition.
- **Bremen Big Data Challenge 2020** 03.2020-04.2020
  Participated in the Big Data Challenge organized by Universität Bremen. The task included predicting the hand movements and actions performed with the help of sensor data (measured with Electromyography (EMG), bipolar derived). Used a combination of U-Net and LSTM models to get an error of 0.464 which was able to secure 12th out of 31 teams.