

Partha Pratim Nath

Machine Learning Engineer | 3D Pose Estimation and Reconstruction

M.Sc. @ TUM | Kurt-Fischer Prize | Vision and AI Research and Startup

Vienna, Austria +49 15239522871 nath.partha@outlook.com nath-partha.github.io [LinkedIn](#)

About Me

- ML Engineer specializing in 3D Vision and Perception Methods with 2+ years designing and refining prototypes from SOTA research
- Most recently built state-of-the-art 3D VLMs for pointcloud-to-layout detection and language-aligned 3D reconstruction from LiDAR scans at scale using A100 clusters.
- Currently seeking roles working on 3D scene/human understanding, robotics, and AR/VR applications in cloud

Key Highlights

- Preparing experiments and team for the Nvidia-DGX-Inception Program(Accepted)
- Implementing Detr fundamentals to pointcloud detection tasks using local attention methods
- Prototyping and Reverse engineering Pointcloud VLMs from SOTA research
- Project Lab Human Activity Understanding creation @ TUM
- Markerless multi-subject hand tracking awarded the Kurt Fischer Prize

Technology Stack

3D Vision & VLMs

PointTransformerV3 Sonata
LLaMA3.2 Qwen2 CLIP
SpatialLM Locate3D

Deep Learning Frameworks

PyTorch Deepspeed Huggingface
PyTorch3D Detectron2 Pointcept
mMLabs

Production ML & Cloud Tools

GCP (A100) AWS (L40, A100)
Docker Multi-node Training
Model Deployment

3D Understanding

Scan2CAD Point Cloud Segmentation
Point Mesh Loss Functions ICP
SLAM Multiview Geometry
Camera Calibration

Programming & Core Tools

Python C++ OpenCV Open3D
NumPy Scikit-learn Pytorch3d
Rerun CVAT

OpenSource Datasets

*ADL4D[1.1M] ScanNet++
Structured3D H2O3D DexYCB
SpatialLM CV4AEC

Technical Reports

ADL4D: Towards A Contextually Rich Dataset for 4D Activities of Daily Living

[Zakour*, Nath*]

- Designed a weakly supervised markerless hand pose annotation method for 4D Human Activity Understanding capable of handling complex multisubject interactions
- Published a Multiview Multisubject activity Dataset of 1.1M frames for complex longhorizon tasks
- Trained and evaluated 3D detection models for human pose, hand mesh recovery, and hand action segmentation
- <https://arxiv.org/pdf/2402.17758>

Multiview Geometry MANO PyTorch3D PyTorch Python

Professional Experience

Circular Pointcloud Analytics GMBH

[Website](#)

Machine Learning and Computer Vision Engineer

Vienna, Austria Dec 2024 - Present

End-to-end R&D and MLOps for large-scale 3D VLM and Scan2BIM pipelines.

ML Ops & Distributed Training Infrastructure

- Scaled pointcept training engine using standalone ZeRO optimisers and model sharding to resolve memory bottlenecks.
- Evaluated DeepSpeed compilation strategies for Pointcept module incompatibilities and workarounds.
- Re-architected training backend by wrapping HuggingFace Trainer, unifying distributed strategies with custom check-pointing and synchronisation for config/source files.
- Deployed on-premise ClearML infrastructure for experiment tracking.
- Evaluated external libraries for unified cloud resource provisioning and training jobs.
- Integrated multi-TB scale datasets for preprocessing and training in segmentation and detection tasks

R&D: 3D Vision-Language & Object Detection

- Integrated LLMs (LLama3.2, Qwen2) with 3D backbones (Ptv3, Sonata) and reverse-engineered SOTA (SpatialLM, Locate3D) to build custom VLM training pipelines on A100 clusters.
- Developed and ablated 3D-Detr/Roomformer architectures. Tested 3D local attention encoders; reformulated losses to improve rotation regression for high-aspect-ratio objects.
- Utilized 150k in GCP startup credits to scale data augmentation and model training experiments on A100 clusters.

Production Engineering & Scan2BIM

- Implemented and containerized Lidar panoptic segmentation pipelines for Scan2BIM/Scan2CAD tasks, deploying robust models (Ptv3, Sonata) for Industry Foundation Classes.
- Maintained production code and managed model updates for out-of-core segmentation pipeline.

Visualization & Strategy

- Technical Communication: Visualized progress using Rerun, Open3D, and high-quality GRUT (Nvidia) renders for internal presentations and the Nvidia-DGX-Inception Program.
- Collaboration: Contributed to core research objectives regarding superpoints, object detection, and 3D reconstruction.

VLM Pointcloud DeepLearning Google Cloud (A100-80x64) Transformers Pytorch3D 3D Detection Deepspeed

RevTec Systems AG: Casinos Austria International

[Website](#)

System Engineer: Machine Learning and Computer Vision

Vienna, Austria Sep 2023 – Sep 2024

Object Detection and Tracking in RGBD images in on-prem Casino surveillance.

- Developed real-time surveillance software tracking currency, gestures, and equipment.
- Built a customer onboarding toolkit using SAM and foundation models to generate customer-specific object models and datasets.
 - Improved legacy code and automated models preparation, successfully reducing installation timelines from 2 weeks to less than 5 days.
 - Evaluated and integrated newer model compilation tools and edge devices for inference scaling options.
- Reviewed and integrated external projects to handle camera calibration and drift stabilization.
- Managed the CVML lifecycle and outreach for alpha customers (UK & ZA), providing rolling updates and leadership demos.

yolov4/8 CVAT SAM Python c++

Infineon Technologies AG

[Website](#)

Software Engineering Intern

Munich, Germany Oct 2022 – May 2023

Developed Machine Vision Software for Human Pose in 3D Camera

- Torch and opensource based detection and tracking | 3D Multiview Calibration.
- Sensor data acquisition library to train gesture detection radar pipelines with cameras.
- Designed scalable calibration routines for multi-camera setups and prototyped RGB-only multiview data acquisition.

Deep Learning PyTorch Computer Vision OOP Windows Git Detectron2 YOLOv8 DVC

Research and Teaching Assistant

Munich, Germany Jun 2021 – Aug 2023

Awarded Kurt Fischer Prize for

- Creating a novel markerless motion capture toolkit [RGB images]
- Creating a very high fidelity human + object interaction dataset that outperformed previous contributions in pose diversity, accuracy and ability to robustly record very long sequences.
- Featured <https://www.ce.cit.tum.de/en/lmt/home/> Slide 6
- Utilised deep learning pose estimation, 3D multiview algorithms and linear mathematical solving to robustly calculate 3D humans in view.
- Benchmark Tasks (3D Tracking, Hand Mesh Recovery, Hand Action Segmentation)

Teaching Assistant: Project Lab Human Activity Understanding (5ECTS)

- Designed Course | Guided Projects | 3DML Topics | ICP . Camera Projection . Rendering | Demo Scripts
- <https://www.ce.cit.tum.de/en/lmt/lehre/projektpraktikum-project-lab-human-activity-understanding/>

Designed Multicamera Studio for RGBD streams with Realsense Sensors in a streamlined Setup

- Low Latency | Extrinsic Calibration | 8-12 Cameras | Distributed ROS | Optitrack Integration

UE4 based VR simulation and photorealistic data capture tool built on <https://sim2realai.github.io/UnrealROX/>

HMR(I2L-Meshnet, HandOccnet)

CNN

X3D

SlowFast

Git

YOLOv8

Weights & Biases

Education

Technical University of Munich

M.Sc. @ TUM School of Computation, Information and Technology

Munich, Bavaria, Germany 2020 – 2023

- CGPA: 1.6
- Kurt-Fischer €1000 Prize

SRM Institute Of Science & Technology

B.Tech. Electronics and Communication Engineering

Chennai, Tamil Nadu, India 2016 – 2020

- First Class with Distinction
- Project: Multispectral Optics Module for a firefighting robot

Languages

English: Native

German: Conversational

Misc: French, Hindi, Bengali

References

Dr Rahul Gopal Chaudhari

TUM Senior Scientist

Maximilian Strobel

Infineon, System Architect Machine Learning

M.Sc. Marsil Zakour

TUM Doctoral Candidate

Michael Winking

Infineon, Staff Engineer