

Dinesh C Dhotrad

Computer Vision Engineer | 3D Vision, SLAM and Sensor Fusion

Dallas, TX | dxhdhotrad.work@icloud.com | +1 (972) 363-3446

linkedin.com/in/dineshdhotrad | github.com/dineshdhotrad | voxaverse.ai | dineshdhotrad.fyi

SUMMARY

Computer vision engineer with 4+ years of industry experience including internships, building 3D reconstruction, photogrammetry, SLAM and sensor-fusion systems in C++ and Python. Shipped a multi-iPhone RGB and LiDAR capture app to the App Store (84 μ s median cross-phone frame sync) and a native C++/Metal 4D Gaussian Splatting engine at 32.97 dB on the N3V benchmark; lead developer of the official Python SDK for Actian VectorAI DB.

SKILLS

3D Vision: 3D reconstruction, photogrammetry, structure from motion, camera calibration, SLAM, ORB-SLAM3, COLMAP, Open3D, 3D Gaussian Splatting, NeRF

Computer Vision & ML: PyTorch, OpenCV, YOLOv8, Mask R-CNN, OneFormer, DINOv2, CLIP, vision-language models, GAN

GPU & Edge: Metal, CUDA, TensorRT, ONNX, NVIDIA Triton, NVIDIA Jetson Orin, NVIDIA Jetson TX2, quantization, PyTorch DDP, pipeline parallelism

NVIDIA Stack: NVIDIA NeMo, NVIDIA VSS, NVIDIA Omniverse, NVIDIA Triton, TensorRT, Jetson

AI Infrastructure: vector databases, Milvus, Qdrant, gRPC, Protocol Buffers, FastAPI, Celery, Redis, PostgreSQL

Languages & Tools: Python, C++, Swift, Rust, C#, Docker

EXPERIENCE

Computer Vision AI/ML Engineer

Nov 2025 – Present

IdeaBoxAI Inc.

Dallas, TX

- Lead developer and release owner of the official Python SDK ([actian-vectorai-client](#)) for Actian VectorAI DB; shipped v1.0.0 to PyPI on the product's launch day; coverage 83% to 99% across 1,628 tests.
- Deployed vision and vision-language model inference with NVIDIA Triton, vLLM, ONNX and TensorRT (FP16/INT8) and multimodal retrieval over Qdrant and Milvus; building NVIDIA NeMo retrieval integrations.
- Benchmarked VectorAI DB against leading open-source vector databases from 50K to 20M vectors with VectorDBBench; [Actian's launch](#) reported more than 22 $\times$ their throughput at 10M vectors.
- Generated SDK code from shared Protocol Buffers definitions with buf and added client-server parity tests that keep the SDK in lockstep with the server API.
- Created SDR-T, a Model Context Protocol server (OAuth 2.1, FastAPI, Celery, Redis, PostgreSQL); fixed a failed customer demo by rewriting tool descriptions: 31/31 correct GPT-4o routing, 491 tests.
- Built the client's production features: sync and async clients over gRPC with REST fallback, TLS, retries with exponential backoff, API-key authentication, a filter DSL and hybrid fusion.

Computer Vision AI/ML Engineer Intern

Sep 2025 – Nov 2025

PranaTree

Tampa, FL

- Developed construction-progress analysis of 360-degree site video on NVIDIA VSS and NeMo with vision-language models, DINOv2/CLIP keyframe extraction and VideoRAG; used NVIDIA Omniverse for digital twins.
- Prototyped robotic site perception with ORB-SLAM3 localization and TensorRT models on NVIDIA Jetson Orin, using Milvus retrieval to match captured site features to plan elements.
- Accelerated 360-degree video ingestion for vision-language model pipelines with an equirectangular-to-cubemap converter (NVENC GPU, CPU fallback).

3D AI Lead Engineer

Jun 2023 – Feb 2024

brahmGAN

Bengaluru, India

- Led core architecture for text-to-3D, image-to-3D and video-to-3D generation (NeRF, photogrammetry) for XR-ready assets.
- Engineered multi-GPU training and inference (PyTorch DDP) with Docker GPU hosting on a decentralized network; 30% lower compute cost in real-world testing.
- Added data masking so personal GPUs on the network could process client workloads privately in exchange for tokens.

Computer Vision Engineer (promoted from Intern and Software Engineer)

Mar 2021 – Apr 2023

Tooliq Innovations LLP

Gurugram, India

- Created a 360-degree video-to-3D pipeline for AEC clients: Mask R-CNN segmentation, OpenMVG/OpenMVS structure from motion and a 3D geometric voting algorithm; 90% background-noise reduction.

- Made UV-texturing projection 15× faster by replacing brute-force nearest-neighbor search with a multithreaded geometric formulation; 10M+ point scans went from hours to minutes.
- Fused RGB, depth and IMU (iPhone LiDAR, Intel RealSense) for SLAM: 25% better trajectory accuracy, 40% faster point-cloud processing with Open3D.
- Engineered the SLAM data layer for an in-house capture device (two RGB and one ToF sensor): RAW depth decoding, stereo point clouds, trajectory generation and loop-closure analysis.
- Deployed real-time 3D capture on NVIDIA Jetson TX2 and iOS (ARKit) through quantization and hardware tuning at 10-15 FPS on device.

Research Project Intern

KLE Technological University

Aug 2020 – Feb 2021

Hubballi, India

- Designed UIE-Net, a GAN for underwater image enhancement for autonomous underwater vehicles: 60-75% visibility improvement; accepted at GC-RDCT 2021.

Project Trainee

Indian Institute of Technology Delhi

Jun 2019 – Jul 2019

New Delhi, India

- With AIIMS Delhi neurosurgeons, built an OpenCV tool that scores micro-suturing performed by trainee neurosurgeons.

PROJECTS

VoxaVerse: 4D Volumetric Video Platform | C++, Metal, Swift, Rust, WebGPU

Jun 2026 – Present

- Shipped [VoxaVerse Beacon on the App Store](#): syncs iPhones into one RGB and LiDAR rig via Kalman-filtered clock sync; 84 μs median frame pairing, 0.389 px camera calibration error.
- Architected a native C++/Metal 4D Gaussian Splatting engine (no CUDA, no Python): 32.97 dB PSNR on the N3V benchmark on a laptop GPU (published state of the art 33.57 dB); 151 FPS mean GPU playback.
- An earlier Apple-native 3D Gaussian Splatting trainer reached 26.22 dB on Mip-NeRF 360 garden in 458 s, above msplat's published 25.68 dB.
- Directed AI coding agents (Claude Code, Codex) across 12+ repos against measured gates; glTF and USDZ exports pass the Khronos validator and Apple usdchecker; [open capture spec](#) with COLMAP import.

Panoptic 3D Point-Cloud Segmentation from 2D Masks | Python, OneFormer, Open3D, RTAB-Map

Jan 2024

- Fused OneFormer panoptic masks (133 categories) from iPhone LiDAR scans into 3D point clouds by multi-view back-projection and voting, with instance splitting and plane refinement.

Vehicle RoadSense ADAS | Python, OpenCV, YOLOv8

Jan 2024

- Real-time vehicle, pedestrian and lane detection combining YOLOv8 with classical lane finding (camera calibration, bird's-eye perspective transform, polynomial fit) at 30+ FPS.

Multi-GPU Pipeline Parallelism for LLMs | PyTorch, CUDA

Aug 2024

- Developed GPipe-style pipeline parallelism for GPT-style transformers on two NVIDIA RTX 2080 Ti GPUs with PyTorch pipeline-parallel APIs: 42% higher training throughput, per-GPU memory from 90% to 65%.

WinPlay: Native Windows AirPlay 2 Sender | C#, .NET 8

Jul 2026

- Implemented AirPlay 2 from scratch ([open source](#): mDNS, HomeKit pairing, RTSP/RAOP, PTP clock for multi-room sync, ALAC, H.264 mirroring); about 850 downloads, with issues filed by external users.

EDUCATION

Master of Science, Computer Science, Case Western Reserve University, Cleveland, OH

May 2025

Coursework: Designing High-Performance Systems for AI, High-Performance Data and Computing, Computer Vision, Digital Image Processing, Robotics, Deep Learning.

Bachelor of Engineering, Computer Science and Engineering, KLE Technological University, Hubballi, India

May 2021

CGPA 8.09/10; member, Center of Excellence in Visual Intelligence (CEVI).

CERTIFICATIONS

- Fundamentals of Accelerated Computing with CUDA C/C++, NVIDIA Deep Learning Institute (2025)
- Developing Secure Software (LFD121), The Linux Foundation (2025)
- Course on 3D Computer Vision, IIIT Hyderabad (2020)
- Triplebyte Certified Data Scientist (2021)

ADDITIONAL INFORMATION

F-1 STEM OPT, authorized to work in the US; will require H-1B sponsorship in the future. Open to relocation anywhere in the US, on-site, hybrid or remote. Languages: English, Kannada, Hindi.