Dhruv Sheth

dsheth@caltech.edu | 626-689-3769 | linkedin.com/in/dhruvmsheth | github.com/dhruvmsheth | dhruvmsheth.github.io

EDUCATION

California Institute of Technology (Caltech) – Bachelor of Science, **Computer Science** **May, 2026 (Expected)**
Coursework: Advanced ML Projects, ML Theory, Algorithms, Software Design, Data Structures, **GPA: 4.0 / 4.0**
Machine Learning and Data Mining*, GPU Programming*, Deep Generative Models*, Large Language and Vision Models*, Computer Systems, Decidability and Tractability, Linear Algebra, Discrete Math, Multivariable Calculus, ODE **ongoing*

WORK EXPERIENCE

Luxonis Corporation | *Software Engineer Intern* **Mar 2021 – Oct 2021**

- Migrated multiple *OpenVINO* quantized models into SDK examples repository by converting them into depthai custom blob format for use with OAK-D boards. Wrote documentation for updates and answered users' queries on forum
- Implemented **EfficientNet-B0** model in *PyTorch* format on OAK-D by writing custom processing script to classify 1000 classes at 28 FPS. Implemented model conversion script from *PyTorch* to intermediate *OpenVINO* to custom depthai blob
- Worked on luxonis' second tangent - commute guardian with **MobileNet SSD** depth-based tracking

EdgeImpulse | *Software Engineer Intern* **Jan 2021 – Dec 2022**

- Beta-tested software before release on *Sony Spresense, Nordic, Himax, Particle*. Wrote extensive documentation for features
- Extended **EdgeImpulse CLI** to Linux and Debian based systems by managing *npm* dependencies and configuring *Node.js* installations. Assisted users on forum with software concerns, debugging and technical queries

RESEARCH EXPERIENCE

Burdick Robotics Group, Caltech | *Summer Undergraduate Research Fellow* **Jun 2024 – Sept 2024**

- Designed **segmentation** module with **perspective warping** using Video Object Segmentation (VOS) network with *PyTorch, TriMesh, Open3D*; **improves SOTA** FoundationPose performance for 6 DoF tracking, recovers from occlusions
- Optimized XMem VOS network using **TensorRT** and gained **20%** speed improvement
- Developed Unity HDRP simulation with C# with **ROS2 TCP/IP** protocol for real-time communication, replicating RealSense D457 with RGBD+IMU sensors. Created additional Unity environment for training **Signed Distance Fields**
- Implemented **Docker** deployment for arm64 (AGX Orin) and x86/64, **CI/CD pipeline** with ROS2. Reduced feature testing time by **83%** with sim2real swapping. Deployed ROS2 package with segmentation and data broadcasting nodes

AMBER Bipedal Robotics Lab, Caltech | *Student Researcher* **Jan 2024 – Mar 2024**

- Improved robustness for a 6D pose estimator for cube tracking in **Unity**: Argus. Research paper in progress
- Engineered a large-scale RGBD data generation pipeline with **MLAgents low-level API** using *PyTorch, MuJoCo, Kornia*
- Finetuned **ResNet-50** using RGBD input from 2 **Zed cameras** improving predictive controller **Cross-Entropy Method (CEM)** performance by **7%** to an average of **30.5** rotations with RGBD compared to **28.5** rotations with RGB

Harvard Medical School, SERI | *Student Researcher at Gang Luo's Lab* **Mar 2022 – Feb 2023**

- Employed a **Faster R-CNN Inception v2 model** on custom eye-gaze dataset to track gazed objects for behavioral understanding on Homonymous Hemianopia (*HH*) patients. Ablation study on YOLO v4 DarkNet, MobileNet v2 SSD
- Conducted large-scale data analysis on 32 minutes of *HH* data using KDE plots, temporal gaze graphs, MF-DFA
- Used a saliency model with **Attentive-ConvLSTM** to estimate deviation of diseased eye-gaze from standard eye-fixations

TECHNICAL SKILLS

Languages: Python, C, C++, C#, Java, TypeScript, MATLAB, SQL/PostgreSQL
Frameworks: PyTorch, TensorFlow, NumPy, Open3D, OpenCV, ROS2, TensorRT, FastAPI, Pinecone, Kaolin, ONNX
Technologies: Docker, CI/CD, Git, APIs, Mujoco, Unity MLAgents, AWS, Linux/Unix, OpenVINO, W&B

PERSONAL PROJECTS

HiveMind (Cursor for teams) –Python, TypeScript (PineCone/FastAPI) – Cerebral Beach Hackathon – *2024*

- AI-powered code editor with context sharing, OpenAI swarm MoE, FastAPI for backend, Pinecone for vector database, and RAG implementation. Built on Continue using Python and TypeScript. Real-time async streaming infrastructure

Jetpack Joyride Clone in C – C, WebAssembly (SDL2/Emscripten) – Software Design Project – *2024*

- Developed JJ clone in C/WebAssembly with file storage, probabilistic difficulty, scrolling background, animation UI

Neural Decoding with sEEG – Python (PyTorch Lightning, SciPy, scikit-learn) – Advanced ML Projects – *2024*

- Developed a neural decoding framework using **A100 GPUs**, extracting 37 visual features via **GPT-4 Vision**.Trained *BiLSTM with Attention*, *LSTM* with grid search achieving **0.69** AUROC and **76%** accuracy on unbalanced classes.

EmotiConnect, Video – Python, Swift (AWS S3, Lambda, Amazon API Gateway) – *BioxML* Hackathon – *2023*

- Implemented a Swift iOS app for Apple Watch to decode emotions using **OpenAI Whisper** for Speech to Text and **AWS Lambda** with GPT-3.5 for processing. Integrated **AWS S3** for storing and *Amazon API Gateway* for streaming to watch