

# Vijay Chevireddi

240-9409294 | [vijaydevmasters@gmail.com](mailto:vijaydevmasters@gmail.com) | [LinkedIn](#) | [GitHub](#) | [Website](#)

## Education

### University of Maryland

*Master of Engineering (M.Eng.) – Robotics* | GPA – 3.77

Coursework: Multi-Modal Models, Natural Language Processing, Deep Learning, Perception, Localization, Path Planning

Aug 2023 – May 2025

College Park, MD

### Osmania University

*Bachelor of Engineering – Mechanical Engineering* | GPA – 9.05/10

Aug 2018 – June 2022

Hyderabad, India

## Skills

**Deep Learning & Machine Learning** – Transformers, CNNs, Object Detection, LLMs, NLP, Multimodal Learning, Generative Models, Fine-Tuning, AI Alignment, distillation, Reinforcement Learning, RLHF

**Cloud and DevOps** – AWS, Amazon SageMaker, Docker, Git

**Programming Languages** – Python, C++, MATLAB, SQL

**Frameworks & Libraries** – PyTorch, TensorFlow, OpenCV, Open3D, Hugging Face Transformers, TIMM, Open AI SDK, LangChain, LangGraph, RAG

**Optimization & Tools** – CUDA, SciPy, NumPy, Pandas, TensorRT, ONNX, Statistical Methods

**Simulations** – CARLA, ROS, Issac Sim, OpenAI Gym, SLAM, Path Planning, Data Visualization

## Experience

### UMD Fischell Department of Bioengineering | Computer Vision Intern

June 2024 – Present

- Implemented YOLOv8 for dredger detection in a University of Maryland initiative to modernize oyster farming, later upgraded to RT-DETR which improved IoU by 6%.
- Developed human pose detection for walking style classification using MMPose and a CNN transformer, enhancing classification accuracy and contributing to improved gait analysis.

### Sai Vamsi Industries | Machine Learning Engineer

Jan 2022 – May 2023

- Integrated a camera-based system using the YOLO framework to detect visual anomalies in press tool machines, showcasing applied computer vision and image understanding skills.
- Combined vibration data from Fluke 3561 FC Vibration Sensors with real-time image analysis to enhance defect detection by 5%, demonstrating practical software development and distributed training methodologies.

## Projects

### Agentic AI Vision-Based Task Planner using DeepSeek's R1-distilled LLaMA 70B Model

Mar 2025 UMD

*Tools and Technologies* – Python, PyTorch, YOLO v8, GroqCloud, OpenCV, CUDA

- Built an autonomous AI workflow using YOLOv8 and DeepSeek's R1-distilled LLaMA-70B to detect objects, predict possible interactions, and plan steps to organize a messy workspace.
- Implemented dynamic task generation with automatic validation and correction, ensuring reliable and executable plans for intelligent scene organization.

### Implementation of GPT and GPT 2 Models from Scratch

Feb 2025 UMD

*Tools and Technologies* - Python, PyTorch, Hugging Face

- Developed a Generative Pretrained Transformer (GPT) model entirely from scratch using PyTorch, implementing core components such as tokenization, self-attention mechanisms, multi-head attention, positional embeddings, and transformer blocks.
- Reproduced a 124M parameter GPT-2 model, including architecture design, model training, fine-tuning, and inference. Optimized training pipelines, configured hyperparameters, and validated model performance against standard language benchmarks.

### Text-to-Command Translation for Robot Navigation Using T5-small Transformer Model | [Link](#)

Dec 2024 UMD

*Tools and Technologies* – Python, PyTorch, Hugging Face, ROS2, Gazebo, CUDA

- Designed a custom transformer-based 3D object detection model for LiDAR point clouds, trained using KITTI dataset.
- Integrated PointNet++ embeddings with custom transformer encoders and designed a novel loss function, enhancing 3D detection and achieved a mAP IoU of 0.67 for 3D bounding box detection on par with existing benchmarks.

### Transformers for 3D Object Detection in LiDAR Point Clouds | [Link](#)

Dec 2024 UMD

*Tools and Technologies* – Python, PyTorch, Open3D, KITTI, CUDA

- Developed and trained a custom transformer-based 3D object detection model for LiDAR point clouds on the KITTI dataset.
- Combined PointNet++ embeddings with specialized transformer encoders and introduced a novel loss function, achieving a 3D bounding box detection mAP IoU of 0.67—on par with leading benchmarks.

### AI learns to Play MARIO Using Deep-Q Learning and SWIN Transformer | [Link](#)

May 2024 UMD

*Tools and Technologies* – Python, PyTorch, Open AI Gym, OpenCV, Git, CUDA

- This project combines DQN with a SWIN Transformer to train Mario-playing AI agent.
- Enabling efficient interpretation of complex visuals and achieving a 500-moving average reward of 2700 in just 1250 episodes—far outperforming a model without SWIN.

### Advanced Vision Systems for Autonomous Navigation Using YOLOv5 and Homography | [Link](#)

May 2024 UMD

*Tools and Technologies* – Python, Pytorch, YOLO v5, OpenCV, ROS2, Gazebo

- Applied homography for dynamic route planning. Integrated optical flow and YOLOv5 to the pipeline to detect obstacles and their velocities in real time, enabling adaptive navigation.
- By leveraging projective geometry to locate vanishing points and horizon lines, the robot achieved over 80% successful navigation.