

SAI MANIKANTA ESWAR MACHARA

☎ +91 9392005772 | ✉ macharasaimanikantaeswar@gmail.com |  [eswarmachara](https://www.linkedin.com/in/eswarmachara)

EDUCATION

RAJIV GANDHI UNIVERSITY OF KNOWLEDGE TECHNOLOGIES, NUZVID IIIT
Pre University Course, CGPA: **9.86**

Nuzvid, Eluru
December 2021 - August 2023

RAJIV GANDHI UNIVERSITY OF KNOWLEDGE TECHNOLOGIES, NUZVID IIIT
Bachelor of Technology in Computer Science, CGPA: **8.95**

Nuzvid, Eluru
August 2023 - May 2027

PUBLICATIONS

S. Retta, **S. M. E. Machara**, I. I. Ganapathi, D. Velayudhan, N. Werghi.
DebrisVision: Bridging the Synthetic-to-Real Gap for Enhanced Underwater Debris Analysis.
In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 2025.

🔗 Google Scholar — CVF Open Access

RESEARCH INTERESTS

My research interests lie primarily in the fields of *Medical Imaging* and *Computer Vision*, and occasionally at their intersection. I am particularly drawn to solving real-world problems through visual understanding and domain-adaptive learning in both scientific and applied contexts.

PROJECTS

AI-POWERED DOCUMENT PROCESSING & VISUALIZATION TOOL

- Led the development of an AI-based prototype for Automated Digitization and evaluation of AP State Police Service Record books at the AI4AP Police Hackathon 2025 (Use Case 6 – Personnel Evaluation).
- Designed a Vision–Language processing pipeline using Qwen2.5-VL for Document Classification and prompt-based structured Field Extraction from unstructured records including promotions, rewards, transfers, complaints, and leave reports.
- Implemented Stamp and Signature localization using custom-trained YOLO models, enabling automated validation of Official Seals and Signatories for audit-ready metadata extraction.
- Built a React–Flask Dashboard for interactive human-in-the-loop verification and storage of validated personnel records in a structured SQLite database.

REFYOLO-HuMAR: MULTI-HUMAN REFERRING EXPRESSION GROUNDING

- Constructed a Human-centric Grounding Dataset of **95K+ images** and **540K+ annotations** by integrating CrowdPose, MPII, OpenImagesV7, and COCO with Automated keypoint–mask generation using ViTPose and SAM2.
- Generated intra-image unique instance-level referring Captions using InternVL 3.5–2B with task-specific prompting for language-guided Human Disambiguation.
- Designed a token-level Text–Visual alignment framework using MiniLM-based query encoding with Bidirectional cross-attention and MIRL-based Hard Negative Mining.
- Achieved **43.23% Margin Success Rate** on a 20K held-out subset; full-scale training and ablation studies in progress for publication.

CIPS-NET: INSTRUCTION-CONDITIONED PATHOLOGY SEGMENTATION

- Proposed a vision-language segmentation framework enabling compositional, instruction-driven segmentation of overlapping pathology classes in histopathology images.
- Introduced a compositional graph reasoning module to explicitly model pathology interactions and inject instruction-aware guidance into dense segmentation decoding.
- Achieved a **Macro Dice score of 0.795** on an instruction-conditioned histopathology benchmark, substantially outperforming fixed-label U-Net baselines.

CARDIOLENS: AI-ASSISTED ECHOCARDIOGRAM ANALYZER

- Developed an AI-powered echocardiogram analysis system using the Stanford's EchoNet Dynamic dataset to automate left ventricle segmentation and ejection fraction calculation and give out detailed health report for patients.
- Worked on Intel DPT Large and ResNet-101 models for segmentation, achieving Dice coefficients of **0.78** and **0.90**, respectively and R2plus1D Model for Ejection Fraction calculation.
- Developed an algorithm to generate electrocardiogram (ECG) waveforms by analyzing heart contractions and expansions using systole and diastole boundaries.

CHEX CLASSIFICATION – MULTIMODAL LEARNING FOR CHEST X-RAYS

- Tackled a multimodal classification challenge in **MAIC** by extracting embeddings from images (Chest X-rays) and text using the OpenAI's CLIP model.
- Developed a custom cross-attention fusion network to effectively combine modalities.
- Achieved a **Macro-averaged F1-score of 0.35**, placing us at **Top 22**, showcasing expertise in Multimodal Deep Learning

THYROID NODULE SEGMENTATION - BENCHMARK STUDY

- Conducted comparative analysis of deep segmentation models including UNet, Attention UNet, TransUNet, U²-Net, and DuckNet.
- Improved segmentation quality and robustness for clinical-grade thyroid nodule detection.

BLoRA-STYLE IMPLICIT TRANSFER LEARNING

- Implemented and analyzed BLoRA on the SubjectPlops dataset, evaluating latent-space knowledge transfer.
- Explored enhancements to Stable Diffusion XL with BLoRA integration for better generative quality.

TECHNICAL SKILLS

Programming Languages: Python, C, Java

Technologies: Machine Learning, Medical Image Analysis, Deep Learning, Computer Vision, LLMs, Generative AI

Tools: VS Code, Google Colab, Kaggle

ACHIEVEMENTS

- **IEEE EMBS BHI 2025 Data Competition Finalist:** Developed MINDER, a machine learning framework for BDI depression score prediction and analysis.
- **2nd Runner-up at IEEE INDICON Intel AI Hackathon 2024, IIT Kharagpur:** Recognized for the **CardioLens project**, an echocardiogram analysis system; placed **3rd** among top research prototypes.
- **22nd Position at MAIC (Multimodal AI4TB Challenge), South Korea, 2024:** Represented **India** at an international AI competition; built a cross-modal model for TB diagnosis from X-rays and clinical notes.
- **Winner at AI4AP Police Hackathon 2025, Guntur:** Developed an AI-powered tool for document analysis and performance visualization; secured **1st place in Use Case 6** (Personnel Evaluation).
- **Top 140, ISRO Bharatiya Antariksha Hackathon 2025:** Selected among national top teams for proposing an innovative hybrid deep learning model crafted specifically for processing high-resolution GIS and remote sensing imagery.