Sai Sri Teja Kuppa

+91-9629609713 kuppasaisriteja@gmail.com linkedin.com/in/saisriteja github.com/saisriteja blog

Education

Indian Institute of Technology, Madras **2023 - 2025**
MS in EE, Specialization in Computer Vision and AI. *CGPA: 8.00/10*

SASTRA Deemed University **2021 Batch**
Bachelor of Technology in Electronics and Communication Engineering *CGPA: 8.08/10*

Research Papers

ICCP (2025, under review) *PhotonSplat: 3D Scene Reconstruction and Colorization from SPAD*
K Sai Sri Teja, Srividya V., Vinayak Gupta, Haejoon, Ashwin Shankaranarayana, Kaushik Mitra.

ECCV Workshop (2024) *DIVA – Deep Indic Virtual Apparel Try-On*
K Sai Sri Teja, Hrishith Mitra, Girish Rongali, Kaushik Mitra.

CVPR Workshop (2024) *MIPI 2024 Challenge on Nighttime Flare Removal: Methods and Results*
Yuekun Dai, ..., **Kuppa Sai Sri Teja**, Jayakar Reddy A, Girish Rongali, Kaushik Mitra.

Multimedia Tools and Applications (2022) *Imaging-Based Cervical Cancer Diagnostics Using Small Object Detection-Generative Adversarial Networks*
Elakkiya R., **K Sai Sri Teja**, Jegatha Deborah L., Bisogni C., Medaglia.

FICTA – Springer (2022) *3D CNN Based Emotion Recognition Using Facial Gestures*
K Sai Sri Teja, Reddy T.V., Sashank M., Revathi A.

Technical Focus Areas

Computer Vision	Image generation (diffusion), image restoration (UNet), 3D vision (Gaussian splitting), vision-language models (LLaMA 3.2, LLaVA), classification (ResNet, ViT), detection (YOLOv5, YOLOv12), novel cameras (SPAD, LiDAR)
Deep Learning	CNNs, transformers, attention mechanisms, encoder-decoder models, contrastive learning, self-supervised learning, quantization.
Systems Engineering	Large-scale video/image processing, Model stress testing, model explainability, model training and pruning, research and module development

Experience

Detect Technologies Pvt Ltd **July 2022 - Jan 2025**
Deep Learning Engineer - II *Chennai, TamilNadu*

- **Project:** Auto Deep Learning Module
  Overview: Developed an automated system for building, training, and testing deep learning models without manual intervention from engineers. The module streamlined data sampling, annotation, model training, and evaluation, significantly improving efficiency and scalability.
  Activities performed:
  - Sampled relevant subsets from large HDD-based datasets based on task-specific criteria, improving workflow efficiency by 50%.
  - Developed an automatic annotation pipeline leveraging foundation models (SAM, CLIP) and in-house trained models, reducing manual annotation effort by 75%.
  - Built a modular framework automating:
    - Data pruning for faster convergence and reduced redundancy.
    - Model training with quantization, halving training time while maintaining accuracy.
    - Model evaluation across classification, detection, and segmentation benchmarks.
    - Generation of explainability reports using Grad-CAM and occlusion-based interpretation.
  - Enabled scaling of incoming projects by 3x through automation.

  Tools: Python, PyTorch, TensorFlow, Linux, Git, WandB, Docker, Captum.

- **Project:** Image Restoration for CCTV Photages
  Overview: Utilized neural networks to clean distorted images by addressing various challenging degradations such as motion blur, low-light, fog, and flare effects. Employed diffusion models and GAN-based architectures to restore and enhance image quality for improved downstream task performance.
  Activities performed:
  * Applied diffusion models (AutoDir) and GANs (Restormer, Uformer) for effective image restoration and enhancement.
  * Tackled diverse degradation types including motion blur, low-light conditions, fog, and flare.
  Tools: PyTorch, AutoDir, Restormer, Uformer, Diffusers
- **Project:** RAG based Gen AI Application
  Overview: Designed and implemented a multi-modal Fishbone Analysis system to support Root Cause Analysis (RCA) and Correction and Protection Analysis (CPA). The system processed client-provided image and text data to deliver actionable insights with enhanced interpretability and accuracy.
  Activities performed:
  * Integrated state-of-the-art Vision-Language Models (LLama 3.2, LLama 3.1, Molmo, LLava) with Large Language Models for robust multi-modal understanding.
  * Improved analysis reliability and interpretability for industrial RCA and CPA workflows using Retrieval-Augmented Generation (RAG) techniques.
  Tools: LLama 3.2, Python, Hugging Face Transformers, Ollama, Huggingface Endpoints, Streamlit.

Healthcare Technology Innovation Center - IIT Madras **Feb 2021 - June 2022**
Project Associate *Chennai, TamilNadu*

- **Project:** Motion Planning for Robotics
  Overview: Designed and implemented a complete software package in Python for robot-assisted needle-based spine surgery. The system includes kinematics, path planning, singularity detection, and collision avoidance. Gained hands-on experience working with real-time robots including KUKA (industrial robot) and Han's Elfin Robot (collaborative robot).
  Activities performed:
  * Built motion planning modules addressing forward/inverse kinematics, collision checks, and trajectory optimization for surgical precision. Developed and deployed robotics algorithms using ROS (Robot Operating System) for real-time execution and control.
  * Integrated camera sensors such as the Intel RealSense camera for depth sensing and 3D scanning in both simulations and real-world scenarios. Conducted extensive simulation experiments using PyBullet and transferred them to physical robots.
  Tools: Python, ROS, PyBullet, Intel RealSense, KUKA, Han's Elfin Robot

Personal Projects

- **Sign Language Person Generation** [GitHub]
  Developed a deep learning pipeline for generating realistic videos of individuals performing sign language from textual or pose-based input.
- **Disfluency Detection in Speech** [GitHub]
  Built a neural network model to detect speech disfluencies (e.g., fillers, repetitions) in spontaneous conversations using NLP and acoustic features.

Achievements

- One among the world top 5 Teams in KUKA Robotics Challenge 2021(Team Aroki)
- Top 5 in the MIPI FLare removal challenge 2024(CVPR 2024).

Certifications and Workshops

- Participated 3D Vision Summer school(3DVSS) at IIIT Banglore(2024).

Extracurricular Activities

Spoken Languages: Telugu(Native), English(Proficient), Tamil(beginner)
Clubs: Member of Robotics Club SASTRA University.
Extra Skills: Handling Peer Pressure, Leadership, Team Work, Research Oriented Vision.