Naresh Kumar Devulapally

Computer Vision, Multimodal AI

devulapa.github.io

About Me

I am an MS Thesis candidate co-advised by Dr. Junsong Yuan and Dr. Sreyasee Das Bhattacharjee at The Visual Computing Lab in the Department of Computer Science at University at Buffalo, SUNY. My Research is focused on Computer Vision and Multimodal AI (Vision-Language models) with applications towards Emotion Recognition, Crowd Analytics, Transformer based Object Understanding, and Emotion Generation. My recent (first-author) works are accepted at ACM Multimedia 2023 (∼25% acceptance rate) and BigMM 2023.

Education

University at Buffalo, SUNY — Buffalo, NY
MS Thesis in Computer Science (Computer Vision) — *August 2022 – May 2024*

National Institute of Technology at Tiruchirappalli — Trichy, TN
B. Tech in Mechanical Engineering with minor in Computer Science — *June 2016 – June 2020*

Selected Publications

- Multi-label Emotion Recognition in Conversation via Multimodal Knowledge Distillation, **Naresh Kumar Devulapally***, Sidharth Anand, Junsong Yuan, Sreyasee Das Bhattacharjee, **ACM Multimedia 2023**.
- AMUSE: Adaptive Multimodal Analysis for Speaker Emotion Recognition in Group Conversations, **Naresh Kumar Devulapally***, Sidharth Anand, Junsong Yuan, Sreyasee Das Bhattacharjee, Yu-Ping Chang, **BigMM 2023**.
- OntoQuest: Automatic Question Generation for online assessments using dynamic ontology-based strategy, **Naresh Kumar Devulapally***, A Santhanavijayan, Deepak G, 15^{th} **ICInPro 2019** (Best Paper Award).

Experience

Graduate Research Assistant — August 2022 – Present
UB Research Foundation — *Buffalo, NY*

- Proposed a self-supervised Multimodal Vision Transformer Framework for Multi-Label Emotion Recognition in videos. Outperformed State-Of-The-Art models by 9% on three large public datasets.
- Proposed a Novel Adaptive Multimodal Fusion and Multi-Head Attention Framework for Speaker Emotion Recognition using video data. Outperformed SoTA by over 7% on two large public datasets.
- Created a large benchmark EmoRace dataset for Emotion Recognition with 250,000 utterances (10 times larger than the current largest benchmark), **Submitted to CVPR 2024**.
- Implemented Fully-Sharded Data Parallel pipeline to train Vision-Transformers across 8 NVIDIA A100 GPUs.
- Led a team of 4 undergrad students at UB to fine-tune Large Language Models for a large medical dataset. Used quantization techniques to improve METEOR and ROUGE score by 10%.

Graduate Teaching Assistant — August 2022 – Present
University at Buffalo, SUNY — *Buffalo, NY*

- Worked as Teaching Assistant to the **Pattern Recognition, Machine Learning, Data Models and Query Languages, Sports Video Analysis using Computer Vision** courses.
- Created interactive lecture slides to elucidate Machine Learning and Computer Vision concepts to over 250 students every semester.
- Published online lecture notes (website format) for **Computer Vision and Image Processing** course.

Lead Data Scientist — Feb 2021 – July 2022
Carbynetech Inc. — *Hyderabad, India*

- Trained a custom Faster R-CNN model, with an ImageNet ResNeXt backbone, to detect P&ID symbols with 94% mAP. Reduced inference time by over 50% as compared to Template Matching.
- Modeled Generative Adversarial Networks (GANs) for synthetic data generation & data balancing.
- Built an Azure Custom Vision AI model to detect and read labels, bar codes, and logos on export pallets with 92% mAP for a multinational engineering and technology client. **Minimized per pallet inspection time by 98%.**
- Increased revenue generated by the Data Science team by over 30%.

- Led end-to-end migration of 800+ GB data from 300+ production facilities worldwide for a multinational client, ensuring 99% migration efficiency.

Machine Learning Engineer July 2020 – January 2021

AIDesign Pvt. Ltd. *Chennai, India*

- Developed Neural Networks to solve real-world fluid flow problems at a faster rate than existing commercial software with less than 5% relative error. Solved run-time & memory bottlenecks.

Projects

Zero-Shot COCO-Google | *HuggingFace CLIP ViT, Flask, Python, Transfer Learning* July 2023 – August 2023

- Implemented text query-based image retrieval web application using OpenAI's CLIP Vision Transformer model. Performed inference optimization for retrieval from 1 Million images in under 25 seconds. Created a HuggingFace space for the project.

AUGMENTify | *Python, Flask, GANs, Image Processing* May 2023 – July 2023

- One-stop solution for Generative Image Augmentation. Implemented a Distributed Training pipeline for text query based Image Processing to style transfer based Generative Image Augmentation.

Neural Network Pruning | *CNNs, Inference Optimization* April 2023 – May 2023

- Optimized VGG16bn model on ImageNet dataset using iterative weight pruning techniques, achieving an accuracy of 98%, model size reduction of 26%, and an inference time speedup of 15%.

Awards and Achievements

- Best Graduate Research Student of the Year Nominee for accepted research work in Computer Vision, Multimodal AI, Real-Time video Emotion Recognition in group conversations.
- Best Paper Award at the Fifteenth International Conference on Information Processing, ICInPro'19.
- First Honorable Mention at Agrusa L. Student Innovation Competition, 2022 for research work on Privacy Preserving Real-Time Emotion Analysis in conversations (PRIMAL).
- Second Place at Agrusa L. Student Innovation Competition, 2023 for leading research work on fine-tuning Large Language Models for Vision and Medical Applications.