

# Sidharth Anand

✉ sidharthanand85@gmail.com | [in](#) LinkedIn | [🌐](#) Website | [🐙](#) GitHub | [📞](#) 8583197724

## RESEARCH EXPERIENCE

---

### Graduate Student Researcher, UC San Diego

Dec 2023 — Ongoing

- Working on a model capable of real-time jazz improvisation in an ensemble based on offline symbolic music generation and online reinforcement-learning audio-based decision making
- Developing a Jazz improvisation dataset that is 5x larger than existing datasets and includes both MIDI and traditionally non-MIDI instruments

### Student Researcher, Visual Computing Lab, SUNY Buffalo

Aug 2023 — Ongoing

- Developed a new state-of-the-art dataset for multimodal emotion recognition and conversation tracking that is 10x larger than existing datasets and is focused on eliminating intrinsic biases present in existing datasets - particularly race and age
- Improve state-of-the-art results on Emotion Recognition in Multi-party conversations by exploring the use of multi-modal data sources, novel fusion and attention techniques, self-supervision and active learning
- Currently focusing on the development of a generalizable framework for improving the robustness of multimodal transformers to lost and corrupted modalities

### Supervised Research, BITS Hyderabad

Jan 2022 — May 2022

- Work on developing lightweight ML and DL models for on-device mobile and IoT malware analysis
- Developed a **CNN** model that achieves performance within 2% of the state-of-the-art while using **300x** fewer parameters and requiring **0.1x** the computation time

## WORK EXPERIENCE

---

### Software Development Intern, Intuit India

May 2022 — Aug 2022

- Worked with Java, SpringBoot, Hibernate and PostgreSQL to bring sweeping changes to the campaign and catalog management system in order to unify disparate data source
- Deployed the respective micro-services using Docker, Kubernetes, Jenkins and ArgoCD
- Rewrote and modernized a critical part of the multi-threading infrastructure for request handling used across the platform resulting in a **15% performance improvement**

### Machine Learning Team Lead, Capital Quant Solutions

Oct 2020 — Jan 2021

- Developed an end-to-end deep learning pipeline for automating document processing workflows
- Led the research and development of custom models for multi-object detection and non-oriented scene text recognition
- The pipeline is currently deployed at the **National Stock Exchange** to automate contract verification and saves hundreds of hours of manual processing

### Visualization Lead and Frontend Engineer, Consilienz

Dec 2019 — Mar 2020

- Developed a versatile system for graph and semantic relationship visualization to aid financial relationship analysis and fraud investigations
- Extend existing JS visualization libraries to develop a visualization system that could scale to 1000s of nodes and tens of thousands of edges while simultaneously be performant and responsive to run on an average laptop browser

## PUBLICATIONS

---

1. **Multi-Label Emotion Analysis in Conversation via Multimodal Knowledge Distillation** in *Proceedings of the 31st ACM International Conference on Multimedia* (2023). <https://doi.org/10.1145/3581783.3612517>.
2. **AMuSE: Adaptive Multimodal Analysis for Speaker Emotion Recognition in Group Conversations.** *Accepted at IEEE Multimedia Big Data 2023.*
3. **A large-scale, diverse dataset for multimodal emotion recognition.** *Under review at CVPR.*
4. **MALITE: Lightweight Malware Detection and Classification for Constrained Devices.** *Under review at IEEE Transactions on Emerging Topics in Computing.*

## PROJECTS

---

### Realistic Playback Generation from MIDI

- Use DL models to generate CC values for raw MIDI and use these values to generate realistic-sounding playback that can rival VSTs
- Drawing parallels from text-to-speech models, explore various transformer architectures and GAN based spectrogram generation to generate audio from MIDI sequences
- Explore various tokenization schemes to represent MIDI

### Software Audio Synthesizer

- Built using C++ and on top of the Windows CoreAudio APIs, the synthesizer supports multiple waveforms, envelopes, LFOs, BiQuad filters, arpeggios and multiple instrument
- Use parallel processing to separate the GUI, Audio calculation and core audio buffer into separate processes. Also used multi-threading to separate each instrument into it's own thread

### PBR Rendering Engine

- Implement a fully-featured implementation of the Unreal4 Physically Based Rendering using C++ and OpenGL
- Supports PBR, HDR, HBAO, Screen-space Reflections, Environment Maps, Area Lights and Temporal Anti-Aliasing all implemented from scratch
- Also uses a tiled-deferred rendering pipeline that is parallel computed on the GPU using compute shaders

### Interlacing for Multi-Path Networking

- Developed a proof-of-concept implementation that integrates interlaced video streams into multi-path networking protocols like MPTCP and MPQUIC using GoLang for MPQUIC and python for MPTCP
- On synthesized and virtual benchmarks shows meaningful improvement over standard streams without noticeable loss in video quality, especially when one of the connection paths is slower than the other

## AWARDS AND HONORS

---

- **Google Summer of Code, 2021** Tempo Overhaul and Smart Tempos for MuseScore
- **Agrusa Student Innovation Challenge, 2022** Honorable Mention for work on Privacy-Preserving and Explainable Emotion Recognition in Conversations

## EDUCATION

---

**Bachelor of Technology, Computer Science** BITS Hyderabad, GPA: 4.00/4.00

2019 — 2023

**Masters of Science, Computer Science** UC San Diego

2023 — Ongoing

## TECHNICAL SKILLS

---

**Languages** : Python, C++, TypeScript, JavaScript, GoLang

**Frameworks** : PyTorch, Tensorflow, NumPy, React, NextJS, Svelte, CUDA, SFML, OpenGL, Vulkan, NodeJS