RAJ PRAKASH GOHIL

✉ rajgohil021@gmail.com +91-9172495348

EDUCATION

Indian Institute of Technology, Kanpur Kanpur, UP
Master of Technology(Research) in Electrical Engineering 2019-2021

Vidyalankar Institute of Technology, Mumbai University Mumbai, India
Bachelors of Technology in Electronics Engineering 2014-2018

WORK EXPERIENCE

Samsung RnD Institute Bangalore India
Senior Machine Learning Engineer August 2022 - Present

- Independent researcher to implement and improve the voice assistant in the day to day life .
- Worked and implemented tasks like Diarization, post stitching of the ASR output, Speaker attributed ASR, ASR Error Correction.
- Algoritms like RPN, Spectral clustering were used for Diarization. Models like BERT were trained for ASR error correction.
- Contributed to various challenges hosted Interspeech, ICASSP, DCASE.
- Major contribution in developing far field source localization.
- Dockerized various algorithm for the usage by the other team members.

Fraunhofer IDMT Germany
Early Stage Researcher in Hearing, Speech and Audio Technology Januray 2022 - August2022

- Implement various state of the art algorithm for speech signal processing.
- Worked on Audio Event Detection and Localization using various directional features.
- Various localization models like SELD, and attention based models are implemented.

Gnani AI India
Speech Processing May 2021 - October 2021

- To develop a system for voice cloning to adapt a unseen speaker using less than 2mins of reference audio.
- Speaker embeddings like xvectors, dvectors and others were studied to use it for speaker adaptation task.
- Various Text-to-Speech models like Tacotron, Fastspeech, Melgan were explored so as to get clean and cloned speech from the text.

Uniphore Software India
Data-Science Intern March 2021 - May 2021

- Analyse the output of the ASR system.
- Develop the model to perform the redaction task for the classified entity from the Hindi text.
- Spacy based NER model was used for training purpose
- Custom dataset was developed by using augmentation techniques.

SONY Research India India
Speech Processing Intern October 2020 - March 2021

- Built a custom data-set for Hindi-English mix audio.
- Build an end to end ASR for Hindi-English with custom acoustic and language model based on Chain models in Kaldi.
- Transformer based LM, CNN based AM is built using and pytorch.
- Various other toolkits like wav2letter, DeepSpeech was also implemented.
- Various pre-processing techniques like noise cancellation, speech activity detection, speaker diarization was also studied and used.
- Tried using various features like Wav2Vec2 in place of MFCC.
- Python wrappers for kaldi like PyKaldi were also studied.
- The WER on the validation data was 13% and the WER on the real dataset(Bollywood-Movies) was 35%.

NMIMS India
Visiting Faculty

- Taught a course on Speech Signal Processing of 60hrs spanned over 4 months.
- Delivered knowledge of speech signals, machine learning, deep learning, speech recognition, how to read research papers.
- The students at the end of the course were able to successfully build a Kaldi based ASR model with in-depth understanding of Language model, acoustic model and their interation.
- Also conducted a viva-voice session where the students presented summary of all the selected speech processing task by efficiently reading the research papers based on the tasks.

Publications

On Combination of various pre-trained models for Speaker Diarization in Multi-Lingual Scenarios*
INTERSPEECH 2023

- JStudy of various pre-trained speaker embedding and vad models.
- Applied novel spectral clustering for speaker diarization.

- Secured second place in the DISPLACE Challenge conducted at Interspeech 2023

Joint DOA Estimation in Spherical Harmonic Domain using Low Complexity CNN
SPCOMM 2022

- Joint DOA estimation is done using low complexity Spherical Harmonics features.
- CNN based learning is adapted so as to capture all information from the features.
- An outlier robust Gross error has been used to measure the accuracy of the DOA estimation.

Learning Based Method for Robust DOA Estimation using Co-prime Circular Conformal Microphone Array
National Conference on Communications(NCC-2021) July 2021

- A direction based features was extracted from audio simulated using Uniform and Co-Prime **conformal** microphone array(C3M) at different SNR for all the source angles.
- Joint estimation of azimuth and elevation has been done using robust direction based GCC-PHAT features.
- A state-of-art method to calculate DOA based on Least Squares was compared with the existing DNN based classifier and with our CNN based classifier.

LEARNING BASED DOA ESTIMATION IN ADVERSE ACOUSTIC ENVIRONMENT USING CO-PRIME CIRCULAR MICROPHONE ARRAY
Asia-Pacific Signal and Information Processing Association December 2020

- A direction based features was extracted from audio simulated using Uniform and Co-Prime microphone array at different SNR for all the source angles(0 to 360).
- A state-of-art method to calculate DOA based on Least Squares was compared with the existing DNN based classifier and with our CNN based classifier.
- Technology and Software used: Matlab, Python, DNN, Multi-Layer perceptron, CNN, Multi-Class Classification, Tensorflow,data generation, MicrophoneArray.

*=Submitted and under review

SKILLS & OTHERS

Data Sciene Library: Pandas, NumPy, SciPy, SciKit-Learn, Matplotlib, ggplot2
Machine Learning: Classification, Regression, Neural Networks, GANs, Autoencoders, Data Visualization
Certifications: Deep Learning AI and Tensorflow and Gudiede projects based on time-series and GAN from Coursera

---

POSITION of RESPONSIBILITY

Overall Placement Cordinator
Conduct the overall placement and internship drive in the campus

- Led a 4-tier team of 150+ members responsible for the placements of 1050 and internship of 750 students.
- Buddy Mentorship Program — Decentralized Work Distribution — Improved Company Slotting Mechanism — Brought internship opportunities for Post-graduates.
- End-to-End online registration of the companies directly on the portal.

Media Head
Transform and store the events in the digital format

- Led a team of 15 members, specialized in photography, videography and editing.
- Produced and directed a short film based on the events of the annual-fest for publicity.
- Organized two photography events and managed 150+ participants from various colleges.

COURSE WORK

Speech Signal Processing	Machine Learning for Signal Processing	Advanced Machine Learning
Statistical Signal Processing	Discrete Signal Processing	Applied Linear Algebra**
Advanced Topics in Machine Learning	Representation and analysis of Random signals	
Natural language Processing	Deep Neural Network**	

**= Online Courses