SARTHAK KUMAR **MAHARANA**

1029 1/2 W 25th St, Los Angeles, CA, USA - 90007

(213)-245-6071 | maharana@usc.edu | https://sarthaxxxxx.github.io/ | sarthaxxxxx

"When something is important enough, you do it even if the odds are not in your favour."

Summary

Master of Science graduate student in Electrical Engineering, with an undergraduate degree in Electrical and Electronics Engineering. Extremely pedantic and hard-working, and very much interested in devising better problem-solving methods for challenging tasks, and learning new technologies and tools if the need arises.

Education

University of Southern California

Los Angeles, USA

Master of Science (MS) *Aug. 2021 - May 2023*

- Major : Electrical Engineering
- GPA : 3.85/4.00
- Coursework : Applications of Machine Learning for Medical Data, Deep Learning Systems, Machine Learning II (Mathematical Foundations and Methods), Machine Learning I (Supervised Methods), Introduction to Digital Signal Processing, Probability for Electrical and Computer Engineers, Linear Algebra for Engineering.

International Institute of Information Technology (IIIT), Bhubaneswar

Bhubaneswar, India

Bachelor of Technology (B.Tech) *Aug. 2016 - Jun. 2020*

- Major : Electrical and Electronics Engineering
- GPA : 8.32/10.00 | Bachelor's Thesis: 9.85/10.00
- Department rank : 6/63

Experience

University of Southern California

Los Angeles, USA

Graduate Research Assistant *May 2022 - Present*

- Currently working at the Laboratory of Neuro Imaging Computing Research (NICR); Advisor: Dr. Yonggang Shi, Associate Professor.
- Developing an end-to-end general pipeline to automate the reconstruction of fiber bundles in the brainstem of the human brain, using diffusion MRI images, supported by an NIH R01 grant.
- Leveraging deep learning based registration and label fusion methods to automatically generate the anatomical ROIs that are critical for fiber bundle reconstruction.

University of Southern California

Los Angeles, USA

Graduate Research Assistant *December 2021 - December 2022*

- Worked at the Signal Analysis and Interpretation Laboratory (SAIL); Advisor: Dr. Shrikanth Narayanan, Professor.
- Performed speaker recognition from rt-MRI videos, based on an unsupervised disentanglement representation learning scheme.
- Contributed to the development of generating embeddings from 2D sagittal-view rt-MRI videos to distinguish between speakers based on their articulatory representations from vocal tract landmarks.

National University of Singapore

Remote

Part-time Research Assistant *July 2020 - April 2021*

- Worked at the Medical Mechatronics Lab , Department of Biomedical Engineering; Advisor: Dr. Ren Hongliang, Associate Professor.
- Experimented with different encoder-decoder architectures (ex. LinkNet) by plugging in spatio-temporal modules (ex. convLSTM) to perform pixel-wise prediction of the needle trajectory in ultrasound images during a kidney biopsy.
- Researched on semantic segmentation, leveraging graph representation learning, in OCT images of a rat colon.
- Proposed the integration of a DGMN (Dynamic Graph Message Passing) network in DGCN (Dual Graph Convolutional Network), for efficient semantic segmentation, to model long-range dependencies in an OCT image.

Indian Institute of Science (IISc) *Bengaluru, India*

Bachelor's Thesis and Student Researcher *December 2019 - September 2020*

- Worked at the Signal Processing and Interpretation (SPIRE) Lab , Department of Electrical Engineering; Advisor: Dr. Prasanta Kumar Ghosh, Associate Professor.
- Studied acoustic-to-articulatory inversion (AAI) model's performance on the dysarthric speech when the model was trained in a corpus-dependent manner using a matched low resource dysarthric corpus or using a mismatched cross-corpus with rich acoustic-articulatory data.
- Investigated the benefit of utilizing cross-corpus acoustic-articulatory data using transfer learning and joint-training techniques for the articulatory predictions of dysarthric subjects.
- Conditioned the jointly-trained AAI model with x-vectors to study its benefits on the AAI performance of dysarthric subjects. Achieved a relative improvement in the Pearson correlation coefficient by $\sim 13.16\%$ and $\sim 16.45\%$ over a randomly-initialized baseline AAI model trained with only dysarthric corpus for dysarthric patients in seen and unseen conditions, respectively. Also achieved relative improvements of $\sim 3.49\%$, $\sim 6.46\%$, and $\sim 4.03\%$ over three baseline AAI models that utilized the cross-corpus, for the seen conditions of dysarthric patients.
- Published first-authored paper at the 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Toronto.

Indian Institute of Technology (IIT), Kharagpur *Kharagpur, India*

Summer Research Intern *May 2019 - July 2019*

- Worked at the Department of Electrical Engineering; Advisor: Dr. Aurobinda Routray, Professor.
- Developed an in-house template matching algorithm, of various phases, to detect breaths in speech recordings using end-to-end deep neural networks.
- Utilised mel-spectrogram coefficients, zero-crossing rate, and the root mean squared energy of varying frame and overlap lengths on the manually segmented breath and non-breath files, for pre-processing.
- Trained the feature matrix on a Convolutional Neural Network (CNN), Long Short Term Memory (LSTM), and Convolutional Recurrent Neural Network (CRNN) and the results were compared for various frame lengths.
- Employed a heuristic technique to join close predicted breath segments, and segments below a certain threshold were removed, for post-processing and to remove any misclassification errors.
- Predicted by fine-tuning different threshold values. The CNN model gave the best F1 score of $\sim 86.14\%$.

National Institute of Technology Karnataka (NITK), Surathkal *Surathkal, India*

Research Assistant *December 2018 - March 2019*

- Worked at the Image Processing and Analysis Laboratory (IPAL), Department of Electronics and Communication Engineering; Advisor: Dr. Shyam Lal, Assistant Professor.
- Developed image processing algorithms to de-haze a single remote-sensing hazy image. Worked on the design and implementation of a dark channel prior approach for de-hazing of images.

Electronics and Radar Development (LRDE), Defense Research Development Organisation (DRDO) *Bengaluru, India*

Summer Intern *May 2018 - July 2018*

- Worked at the Radar I Division; Advisor: Mr. Paramananda Jena, Scientist 'F'.
- Trained on FPGAs, digital designs using VHDL , and signal processing by making use of ModelSim and Altera Quartus II, as base platforms.
- Designed a low pass FIR filter for ECG denoising to remove the baseline wander, EMG interference and power line noise.
- Implemented an IEEE paper titled 'Optimal Factoring of FIR Filters'.

Teaching

USC Ming Hsieh Department of Electrical and Computer Engineering *Los Angeles, USA*

Course Grader *January 2022 - May 2022*

- Worked as a mentor/grader for a graduate-level course, EE 541: A Computational Introduction to Deep Learning.
- Responsibilities included the grading of assignments, holding office hours, monitoring online forums, and project grading.

Publications

Published

- [C2] **Sarthak Kumar Maharana**, Aravind Illa, Renuka Mannem, Yamini Belur, Preetie Shetty, Veeramani Preethish Kumar, Seena Vengalil, Kiran Polavarapu, Nalini Atchayaram, and Prasanta Kumar Ghosh, "Acoustic-to-articulatory inversion of dysarthric speech by using cross-corpus acoustic-articulatory data," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2021, pp. 6458-6462. [LINK]
- [C1] Jayanta kumar Sahu, Sudhakar Sahu, JP Patra, **Sarthak Kumar Maharana**, and Bhagabat Panda, "Harmonics analysis of a PV integrated hysteresis current control inverter connected with grid and without grid," in *2019 International Conference on Smart Systems and Inventive Technology (ICSSIT)*. IEEE, 2019, pp. 1154–1157. [LINK]

Skills

- **Languages:** Python, MATLAB, Bash
- **Libraries:** PyTorch, Keras, TensorFlow, OpenCV, scikit-learn
- **Others:** Git, LaTeX
- **OS:** Unix, Windows

Projects

Acoustic-to-Articulatory Inversion of Dysarthric Speech using Self-Supervised Learning [LINK]

Los Angeles, USA

January 2023 - May 2023

- First work that demonstrated the effects of features from pre-trained self-supervised learning models, for acoustic-to-articulatory inversion (AAI) of dysarthric speech.
- Experimented by conditioning the BLSTM AAI model with x-vectors i.e., speaker-specific embeddings. Results revealed that DeCoAR, in the fine-tuned scheme, achieved a relative improvement of the Pearson Correlation Coefficient (CC) by $\sim 1.81\%$ and $\sim 4.56\%$ for healthy controls and patients, respectively, over MFCCs.
- Concluded that SSL models, like wav2vec, APC, and DeCoAR, that are trained with feature reconstruction or future timestep prediction tasks, perform well across all three AAI training schemes for predicting dysarthric articulatory trajectories.

Understanding Multi-Modal Speaker Recognition via Disentangled Representation Learning [LINK]

Los Angeles, USA

November 2022 - December 2022

- Presented an adversarial invariance approach to address the problem of multi-modal speaker recognition, by learning two split representations of the input that were robust to various sources of variability present in videos and speech.
- Proposed a 2DCNN and LSTM autoencoder + Global Average Pooling (GAP) based feature extraction pipeline to extract robust frame-level features.
- Experimental results revealed that the multimodal learning setup achieved a relative improvement of $\sim 25.5\%$ in classification accuracy over using x-vectors only and about $\sim 0.048\%$ over using the frame-level features only, as the input.
- Achieved even superior results when the video frames were perturbed with Gaussian noise, which made the adversarial invariance network more robust.

Understanding Linguistic Patterns for Text-Based Speaker Classification [LINK]

Los Angeles, USA

November 2022 - December 2022

- Studied various text feature extraction methodologies using pre-trained models and classification algorithms and compared them to build a computationally efficient system targeted for text-based speaker classification.
- Performed binary classification - predict if a dialogue had been said by one of the three main characters against any other character in the show (South Park). Multiclass classification - identify which of the three main characters had spoken a dialogue.
- SVM with the use of SBERT embeddings, as features, achieved $\sim 68.2\%$ for the binary problem and $\sim 50.4\%$ for the multi-class problem, in terms of classification accuracy. It achieved an overall improvement of $\sim 10\%$ over a transfer learning baseline system (SVM classifier) using word2vec text embeddings.
- Results suggested that long-term dependencies among words and contextual information were crucial while doing speaker recognition/classification in a large corpus.

The Effect of Conditioning of Trigonometric Transformations of Dates with Meteorological Data in Forest Fires Prediction: An Experimental Study [LINK]

Los Angeles, USA

April 2022 - May 2022

- Studied the effects of conditioning a trigonometric transformation of dates with meteorological data, that would aid in predicting the occurrence of forest fires in Algeria.
- Computed statistical parameters (mean/maximum/minimum/median) of important weather variables.
- Conditioned the cosine of the days with the best features, after feature engineering.
- SVM with a linear kernel and a regularisation parameter of value 5, without any conditioning, gave the best results with a test accuracy of 91.6% and F-1 score of 89.3%, and a relative improvement of $\sim 5.84\%$ in accuracy and $\sim 7.2\%$ in F1 score, over a nearest-means baseline classifier. The corresponding SVM model conditioned with the transformed dates gave a relative improvement of $\sim 3.92\%$ in classification accuracy, over the baseline.
- The course project received the highest grade, in a class of about 250 students.

Sentiment Analysis : Zomato, Bengaluru [LINK]

Bhubaneswar, India

October 2019

- Objectives: To understand the popularity of foods in a locality and the sentiments of customer reviews, distribution of restaurants, the approximate price of foods based on cuisines, and classify review comments as positive, negative, or neutral.
- Performed exploratory data analysis to understand the database behaviour, generated word clouds of cuisines liked by people, and extracted locations using GeoPy.
- Topic modeling of positive and negative comments received by customers; t-SNE, for visualization, of adjectives used in such comments.
- Deployed an LSTM model for the sentimental analysis of customer comments based on keywords and adjectives; received an accuracy of 0.954 on the test set.

FreeSound General-Purpose Audio Tagging Challenge [LINK] *Bhubaneswar, India*

July 2019

- Objective: Classify audio signals into 10 different classes.
- Dataset description: 10 classes of sounds of different musical instruments, each having 30 audio files.
- The input features were trained on both CNN and a basic RNN for comparison, using Keras as the deep learning framework. Features : MFCC (Mel-Frequency Cepstral Coefficents).
- CNN gave a better accuracy of about 96%. Predictions were made on the training set itself to mark the probabilities.

Analysis and Implementation of Clustering Algorithms *Bhubaneswar, India*

March 2019

- Implemented K-Nearest Neighbors (KNN) and K-Means in Python and MATLAB to facilitate visualization of clustering.
- Designed a real-time face recognition system using KNN.
- Designed an image segmentation system using K-Means.

Single Image Haze Removal using Dark Channel Prior [LINK] *Bhubaneswar, India*

January 2019

- The dark channel is based on the following observation relating to outdoor images: In most of the non-sky patches, atleast a certain colour channel has at least one dark channel i.e it has low intensities tending to zero. With this observation in mind, an algorithm was used for the haze removal of remote sensing images in MATLAB.

Accomplishments

2020	**Governing Body Merit Scholarship,** Academic year: 2019-2020, IIIT-Bh	*Bhubaneswar, India*
2019	**Dean's List,** 3^{rd} year, IIIT-Bh [LINK]	*Bhubaneswar, India*
2019	**Summer Research Fellowship (SRF) 2019,** The Indian Academy of Sciences (< 10% selection rate) [LINK]	*Bhubaneswar, India*
2016	**95 percentile among 1.3 million students,** Joint Entrance Examination (JEE) Main, 2016.	*Bengaluru, India*
2014	10^{th} **Board Exam scholarship,** Rank 5 among a graduating class of 120 students.	*Bengaluru, India*
2011	**All India Rank (AIR) of 800,** National Science Talent Search Examination (NSTSE)-2011.	*Bengaluru, India*

Activities

USC IEEE Graduate Society *Los Angeles, USA*

Member *October 2021 - Present*

- Member of USC IEEE's Graduate student chapter.
- Responsibilities: Attend group meetings, strengthen academic and social growth of the members, and host workshops.

BangPypers - The Bangalore Python Usergroup *Bengaluru, India*

Member *September 2020 - Present*

- BangPypers, the Bangalore Python Usergroup is one of the largest and oldest Python user groups in India.
- Part of a group, that on a weekly basis, discusses python-related topics.

PyCon India 2020 *Bengaluru, India*

Member of the Social Media Team *August 2020 - October 2020*

- PyCon India is a conference, and the largest gathering, of Pythonistas in India for the Python programming language.
- Responsibilities: content writer for social media handles, helped the promotions team to reach out to organisations and colleges, and interacted with individuals who have contributed to the language, and also worked on creating virtual swags.

International Association of Engineers (IAENG) *Kowloon, Hong Kong*

Student Member *July 2018 - Present*

- Student member of IAENG Society of Electrical Engineering.
- Access to view all IAENG journal and conference publications free of cost, and attend meetings.