Saurabh Kumar

Google Scholar: tinyurl.com/4pyub4pd
GitHub: github.com/saurabhk0317
Email: saurabhk0317@gmail.com
LinkedIn: saurabhk0317

Education

B.Tech. in Electronics and Communication Engineering 2016–2020
National Institute of Technology, Patna
GPA: 7.42/10.00

- **Capstone Project:** "Simulation of Hybrid Plasmonic Waveguides Based Devices and Its Applications"
  - **Description:** Conducted simulations and analyzed the characteristics of various hybrid plasmonic waveguide-based devices using COMSOL Multiphysics software.
  - **Advisor:** Dr. R. Ranjan

Experience

Research Associate July 2021–Present
SPIRE Lab, IISc Bangalore, India
Advisor: Prof. Prasanta Kumar Ghosh
Projects:

- *Speech Recognition in Agriculture and Finance for the Poor in India (RESPIN)*
  Publications: [2], [3]
  - Primarily involved in developing automatic speech recognition (ASR) models and resources in Indian languages.
  - Exploring the effect of auxiliary tasks such as dialect and domain identification on ASR.
  - Conducted a comprehensive study on the impact of dialect and domain-specific statistical language models on ASR in Indian languages.
- *Text-to-speech synthesizer in nine Indian languages (SYSPIN)*
  Publications: [1]
  - Developed an efficient pipeline using Kaldi-based ASR models to filter out poorly aligned/transcribed data, contributing to the quality assessment of the SYSPIN data.

Machine Learning/Signal Processing Consulting Engineer September 2022–Present
ARTPARK, IISc Bangalore, India
Project: Development and Curation of Speech and Text Corpora for VAANI

- Playing a pivotal role in developing and curating extensive speech and text corpora covering 773 districts in India for the VAANI project, funded by Google.
- Employing signal processing and machine learning techniques to create an efficient data pipeline, ensuring data quality and accuracy in the VAANI project.
- Contributing to the success of the VAANI initiative by maintaining the highest standards of data quality.

Undergraduate Student Researcher July 2019–June 2020, December 2020–June 2021
Department of ECE, NIT Patna, India
Advisor: Dr. S. Shahanawazuddin
Publications: [5], [6]

- Explored front-end acoustic feature enhancement techniques to improve children's speech recognition in noisy and low-resource conditions [6].
- Developed robust children's speech recognition systems in low-resource conditions and investigated the effect of formant scaling-based data augmentation on children's speaker verification systems [5].

Summer Intern May 2019–June 2019
EICT Academy, IIT Guwahati, India
Advisor: Dr. Gaurav Trivedi
Project: Cascade filter bank implementation of discrete wavelet transform

- Designed Verilog code for a 5-stage cascade filter bank structure of the discrete wavelet transform for FPGA-based simulation, gaining valuable insights into FPGA simulation.

Conferences and Workshops

LIMMITS'24 Challenge - ICASSP 2024

- Co-hosted a special session on *LIMMITS'24: Multi-speaker, Multi-lingual Indic TTS with voice cloning* at ICASSP 2024 in Seoul, Korea.
- Contributed to the data curation and the development of baseline models released through the challenge.

MADASR Challenge - ASRU 2023 Workshop

- Organizing team member for the MADASR (Model Adaptation for ASR in low-resource Indian languages) challenge at the ASRU 2023 workshop.
- Contributed to the curation of training and testing data, as well as the development of baseline ASR models and recipes in Bengali and Bhojpuri.

SLT-CODE Hackathon 2022

- Award-Winning Project: "Dialectical speech recognition for two Indian languages - Bengali and Bhojpuri."
- Our project received the Best Hackathon Project Award at SLT-CODE Hackathon 2022.
- Contributed to curating training and testing data and developed Kaldi-based ASR models for Bengali and Bhojpuri.
- Presented our project's findings at SLT 2022.

Gram Vaani ASR Challenge - Interspeech 2022

- Organizing team member for the Gram Vaani ASR challenge at Interspeech 2022.
- Contributed to curating 1100 hours of Hindi ASR corpus, open-sourced to the research community through the Gram Vaani challenge [4].
- Developed and open-sourced Kaldi-based baseline ASR models and recipes.

Publications

[1] A. Singh, A. Jayakumar, Deekshitha G, H. Tiwari, J. Bandekar, S. Badiger, S. Udupa, **Saurabh Kumar**, and P. K. Ghosh, "An End-to-End TTS Model in Chhattisgarhi, a Low-Resource Indian Language", in *Speech and Computer*, Cham: Springer Nature Switzerland, 2023, pp. 164–172, ISBN: 978-3-031-48312-7.

[2] A. Singh, A. S. Mehta, K. S. Ashish Khuraishi, G. Deekshitha, G. Date, J. Nanavati, J. Bandekar, K. Basumatary, P. Karthika, S. Badiger, S. Udupa, **Saurabh Kumar**, P. K. Ghosh, V. Prashanthi, P. Pai, R. Nanavati, S. P. R. Mora, and S. Raghavan, "An ASR Corpus in Chhattisgarhi, a Low Resource Indian Language", in *Speech and Computer*, Cham: Springer Nature Switzerland, 2023, pp. 173–181, ISBN: 978-3-031-48312-7.

[3] S. Udupa, J. Bandekar, G. Deekshitha, **Saurabh Kumar**, P. K. Ghosh, S. Badiger, A. Singh, S. Murthy, P. Pai, S. Raghavan, and R. Nanavati, "Gated Multi Encoders and Multitask Objectives for Dialectal Speech Recognition in Indian Languages", in *2023 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU)*, 2023, pp. 1–8.

[4] A. Bhanushali, G. Bridgman, D. G, P. Ghosh, P. Kumar, **Saurabh Kumar**, A. Raj Kolladath, N. Ravi, A. Seth, A. Seth, A. Singh, V. Sukhadia, U. S, S. Udupa, and L. V. S. V. D. Prasad, "Gram Vaani ASR Challenge on spontaneous telephone speech recordings in regional variations of Hindi", in *Proc. Interspeech 2022*, 2022, pp. 3548–3552.

[5] S. Shahnawazuddin, A. Kumar, V. Kumar, **Saurabh Kumar**, and W. Ahmad, "Robust children's speech recognition in zero resource condition", *Applied Acoustics*, vol. 185, p. 108 382, 2022, ISSN: 0003-682X.

[6] S. Shahnawazuddin, A. Kumar, **Saurabh Kumar**, and W. Ahmad, "Enhancing robustness of zero resource children's speech recognition system through bispectrum based front-end acoustic features", *Digital Signal Processing*, vol. 118, p. 103 226, 2021, ISSN: 1051-2004.

Skills

- **ML/DL Toolkits:** PyTorch, Scikit-learn
- **ASR Frameworks:** ESPnet, Kaldi, Fairseq
- **Others:** MATLAB, Audacity, Praat, Git

Languages

- **Programming:** Python, Bash, PHP
- **Natural:** Bhojpuri, English, Hindi, Maithili
  - **TOEFL:** 100/120

Extracurricular Activities

- Served as a student volunteer at ICASSP 2024, assisting with various organizational tasks, guide assistance and session management
- Web coordinator of Yoga and Meditation Club of NIT Patna in 2018-19.
- Member of the organizing committee of the annual cultural fest of NIT Patna in 2018.
- Co-organized a fundraising project to help flood-affected people of Bihar in 2017.