

Soumya Dutta

Research Scientist, Adobe Research

📞 (+91)8828290596

✉ soumyadutta@iisc.ac.in

📄 <https://soumya-dutta.github.io>

🔍 Google scholar: <https://tinyurl.com/2tv78cfp>

Research Interests

- Multimodal Document Transformation
- Multimodal Emotion Recognition
- Synthesis of Emotional Speech
- Adapting LLMs/multimodal LLMs for emotion recognition

Education

- 2021-2026 **Indian Institute of Science Bangalore.**
Ph.D in Electrical Engineering **CPI: 10/10**
- 2018 **Indian Institute of Technology Bombay.**
M.Tech in Control and Computing **CPI: 9.41/10**
- 2015 **Indian Institute of Engineering Science and Technology Shibpur.**
B.E. in Electrical Engineering **CPI: 8.99/10**

Publications

Journals

- **S. Dutta**, S. Ganapathy, “**Leveraging Content and Acoustic Representations for Speech Emotion Recognition**” *IEEE Transactions on Audio, Speech, and Language Processing*. [Paper] [Code]
- **S.Dutta**, S.Balaji, S.Ganapathy “**A Mixture-of-Experts Model for Multimodal Emotion Recognition in Conversations**” *Computer Speech & Language* [Paper] [Code]
- **S.Dutta**, A. Jain, S.Ganapathy, “**Textless and Non-Parallel Speech-to-Speech Emotion Style Transfer**” *IEEE Transactions on Audio, Speech, and Language Processing (Under Review)* [Paper] [Code]

Conferences

- **S.Dutta**, S.Balaji, Varada R, V.Salinamakki, S.Ganapathy “**ABHINAYA - A System for Speech Emotion Recognition In Naturalistic Conditions Challenge**” *Interspeech 2025* [Paper] [Code]
- **S.Dutta**, S.Ganapathy, “**LLM supervised Pre-training for Multimodal Emotion Recognition in Conversations**” *ICASSP 2025* [Paper]
- **S.Dutta**, S.Ganapathy, “**Zero Shot Audio to Audio Emotion Transfer With Speaker Disentanglement**” *ICASSP 2024* [Paper] [Code]
- **S.Dutta**, S.Ganapathy, “**Multimodal Transformer With Learnable Frontend And Self Attention For Emotion Recognition**”, *ICASSP 2022* [Paper] [Code]
- **S.Dutta**, N.Rangaraj, M.N.Belur, S.Dangayach and K.N.Singh, “**Construction of periodic timetables on a suburban rail network-case study from Mumbai**”, RailLille 2017

Work Experience

- May 2025 - **Position: PhD Research Intern, Adobe Research Bangalore.**
- Aug 2025
 - Worked on the problem of generating **faithful podcast transcripts** from documents
 - Proposed a dataset and evaluated faithfulness for script generation from open-source and closed-source LLMs
 - Proposed a model-agnostic mitigation strategy for improving the faithfulness of the generated scripts

- Jan 2024 - **Position: PhD Research Intern, Samsung Research Institute Bangalore.**
- Jul 2024
- Worked on the problem of **textless speech-to-speech translation** system for French-English, Spanish-English and German-English language pairs
 - Developed a baseline system inspired by an existing work called TranSentence for this task
 - Implemented **expressive speech translation** on top of the baseline system
- Jul 2018 - **Position: Cognitive Data Scientist.**
- Feb 2021 Organization: **IBM**
- Worked on a **SVM Classifier** for email intent classification with a **precision of 76%** and **recall of 91%**
 - Worked on a **Virtual Makeup Try-On** system with lips and hair segmentation followed by color transfer from example lipstick and hair-dye patches to lips and hair respectively. The color transfer was done by **matching the distribution** of the source and target.
 - Detected human beings in a video by using a **YOLOv5m network pre-trained on COCO dataset**. This was used along with a **tracking algorithm (SORT)** to raise an alert if a person was loitering in an area

Scholastic Achievements

- **Qualcomm Innovation Fellowship 2025 Winner**-Multimedia Computing Track
- Awarded the best presentation award at **EECS Symposium IISc 2025**
- **Qualcomm Innovation Fellowship 2023 Finalist** (awarded to 37 out of 124 teams)
- Recipient of the **Prime Minister's Research Fellowship(PMRF)**, 2022
- Awarded the **SPS travel grant** for attending ICASSP 2022 (awarded to 56 out of 3967 papers)
- Secured a rank of **231** in **WBJEE 2011** out of approximately *1,30,000 students*

Skills

Programming Python, C, C++, Pytorch
 Software MATLAB, L^AT_EX, MS-Office

Relevant Course Work

Ph.D Machine Learning for Signal Processing, Advanced Deep Learning, Stochastic Models and Applications, Optimization for Machine Learning and Data Science, Detection and Estimation Theory, Speech Information Processing

Graduate Machine Learning, Applied Linear Algebra, Statistical Signal Analysis, Optimization Models