SAURABH KATARIA, Ph.D.

saurabhkataria10@gmail.com (preferred) ⋄ saurabh.kataria@emory.edu ⋄ LinkedIn
saurabh-kataria.github.io (Personal website) ⋄ Google Scholar

RESEARCH INTERESTS

Robust Audio processing, Paralinguistics, Foundation models, Machine Learning

EXPERIENCE

Postdoctoral Fellow, *Center for Data Science*, *School of Nursing, Emory University* — Dec'23 - present
Mentor: Prof. Xiao Hu
Photoplethysmography-sensor based foundation models for human health assessment

EDUCATION

Doctor of Philosophy (Ph.D.), Electrical Engineering, *Johns Hopkins University*, USA — Aug'17 - Jul'23
Thesis: "Robust Speaker Recognition using Perceptual and Adversarial Speech Enhancement" [link]
Advisors: Prof. Jesús Villalba, Prof. Najim Dehak
Affil: *Center for Speech and Language Processing* (CLSP), *Human Language Technology Center of Excellence* (HLTCOE)

Master of Science in Engineering (M.S.E.), *Johns Hopkins University*, USA — Aug'17 - May'20
Electrical and Computer Engineering (obtained while pursuing Ph.D.) — GPA: 3.8/4

B.Tech. - M.Tech. dual degree program, *Indian Institute of Technology (IIT) Kanpur*, India — Jul'12 - Jun'17
- **Master of Technology (M.Tech.)**, Electrical Engineering — GPA: 9.0/10

Thesis: "Unsupervised and Supervised Saliency Estimation Methods for Affective Speech"
Advisors: Prof. Tanaya Guha, Prof. Rajesh Hegde
- **Bachelor of Technology (B.Tech.)**, Electrical Engineering — GPA: 8.4/10
- **Minor in Artificial Intelligence**, Computer Science and Engineering

INTERNSHIPS

Research Intern, *Tencent America* — May'20 - Aug'20
Mentor: Dr. Shi-Xiong (Austin) Zhang, Leader: Dr. Dong Yu — Seattle, WA, USA (remote)
Multi-channel Audio-Visual speech enhancement and speaker recognition [link]

Student Researcher, *Frederick Jelinek Memorial Summer Workshop (JSALT)* [link] — May'19 - Aug'19
Venue: École de Technologie Supérieure (ÉTS), Organizer: CLSP, Johns Hopkins University — Montreal, Canada
Mentors: Prof. Jesus Villalba, Prof. Najim Dehak, Leader: Dr. Paola Garcia
Speech enhancement using auxiliary DNNs [link]

Research Intern, *INRIA* — May'16 - Aug'16
Mentor: Dr. Antoine Deleforge, Leader: Dr. Rémi Gribonval (Team PANAMA) — Rennes, France
Estimating room acoustics and data simulation [link]

Research Intern, *New York University* — May'15 - Aug'15
Mentor: Prof. Siddharth Garg — New York City, NY, USA
De-anonymizing social network graphs

POSITIONS

- Junior researcher, Editorial Board, Physiological Measurement (PMEA) journal, 2025-present
- Full invited member, Sigma Xi: The Scientific Research Honor Society, 2025
- Expert peer reviewer for *BMJ Digital Health & AI* 25-26

NEWS ARTICLES

- SIAM - https://www.siam.org/publications/siam-news/articles/foundation-models-for-physiological-data/

AWARDS

- Emory University Postdoctoral Association Professional Development Travel Award 2025 (research merit-focused)
- Johns Hopkins University ECE department **graduate fellowship** (2017-2018) (extended to 1.5 yrs)
- **Best project award** in the course EE698M (Topics in Image and Signal Processing; Prof. Tanaya Guha), 2016 titled "Scene intensity estimation and ranking for movie scenes through direct content analysis"
- Awarded **two travel grants** by Dept. of Science and Technology (DST), Govt. of India (2015) and IIT Kanpur (EE).

ORGANIZER

1. Symposium with University of Bologna | IEEE EMBC'25 | Foundation models and Pain management in Oncology
2. 8th Mid-Atlantic Student Colloquium on Speech, Language and Learning (MASC-SLL 2020), Program Organizing Committee, University of Maryland, College Park

FIRST-AUTHOR PUBLICATIONS

1. **Kataria, S.**, Xiao, R., Ruchti, T., Clark, M., Lu, J., Lee, R.J., Grunwell, J. and Hu, X., 2025, July. *Continuous cardiac arrest prediction in icu using ppg foundation model.* In 2025 47th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (pp. 1-6). IEEE. [pdf] Oral
2. **Kataria, S.**, Villalba, J., Moro-Velázquez, L., Żelasko, P. and Dehak, N., *Time-domain speech super-resolution with GAN based modeling for telephony speaker verification.* IEEE/ACM Transactions of Speech, Audio, and Language journal 2024 [pdf]
3. **Kataria, S.**, Villalba, J., Moro-Velázquez, L., Thebaud, T. and Dehak, N., 2023. *Self-FiLM: Conditioning GANs with self-supervised representations for bandwidth extension based speaker recognition* INTERSPEECH 2023 [pdf] (Oral)
4. **Kataria, S.**, Villalba, J., Moro-Velázquez, L. and Dehak, N., *Joint domain adaptation and speech bandwidth extension using time-domain GANs for speaker verification.* INTERSPEECH 2022 [pdf] (Oral)
5. Joshi, S.*, **Kataria, S.***, Villalba, J. and Dehak, N., *AdvEst: Adversarial Perturbation Estimation to Classify and Detect Adversarial Attacks against Speaker Identification.* INTERSPEECH 2022 [pdf] (Special Session)
6. **Kataria, S.**, Villalba, J. and Dehak, N., *Perceptual loss based speech denoising with an ensemble of audio pattern recognition and self-supervised models.* ICASSP 2021 [pdf]
7. **Kataria, S.**, Villalba, J., Żelasko, P., Moro-Velázquez, L. and Dehak, N., *Deep feature cyclegans: Speaker identity preserving non-parallel microphone-telephone domain adaptation for speaker verification.* INTERSPEECH 2021 [pdf]
8. **Kataria, S.***, Zhang, S.X.* and Yu, D., *Multi-Channel Speaker Verification for Single and Multi-Talker Speech.* INTERSPEECH 2021 [pdf]
9. **Kataria, S.**, Nidadavolu, P.S., Villalba, J. and Dehak, N., *Analysis of deep feature loss based enhancement for speaker verification.* Odyssey 2020 [pdf]
10. **Kataria, S.**, Nidadavolu, P.S., Villalba, J., Chen, N., Garcia-Perera, P. and Dehak, N., *Feature enhancement with deep feature losses for speaker verification.* ICASSP 2020 [pdf]
11. **Kataria, S.**, Gaultier, C. and Deleforge, A, *Hearing in a shoe-box: binaural source position and wall absorption estimation using virtually supervised learning.* ICASSP 2017 [pdf] (Oral)

OTHER PUBLICATIONS

1. Panchumarthi, L. Y., Meng, Z., Lu, A., Kim, T., **Kataria, S.**, Fedorov, A., Bold, D., Gow, B., Pollard, T. J., Zègre-Hemsey, J., Dzikowicz, D., Hu, X., van Assen, M., Wright, D. W., Kumar, L., & Xiao, R. CardioFM: A multimodal foundation model for ECG representation learning using waveform reconstruction and clinical text alignment. 2026 International Society for Computerized Electrocardiology (ISCE) Conference.
2. Ni, J., **Kataria, S.**, Tang, S., Yang, C., Hu, X. and Jin, W., 2025. *Ppg-distill: Efficient photoplethysmography signals analysis via foundation model distillation.* (NeurIPS 2025 TS4H workshop Spotlight) [pdf]

3. Wang, M., Yan, R., Li, C., **Kataria, S.**, Hu, X., Clark, M., Ruchti, T., Buchman, T.G., Bhavani, S.V. and Lee, R.J., 2025. *Estimating Clinical Lab Test Result Trajectories from PPG using Physiological Foundation Model and Patient-Aware State Space Model–a UNIPHY+ Approach.* NeurIPS 2025 TS4H workshop [pdf]

4. Chen, Z., Ding, C., **Kataria, S.**, Yan, R., Wang, M., Lee, R. and Hu, X., 2025. *GPT-PPG: a GPT-based foundation model for photoplethysmography signals.* Physiological Measurement, 46(5), p.055004. [pdf]

5. Meng, Z., Panchumarthi, L.Y., **Kataria, S.**, Fedorov, A., Zègre-Hemsey, J., Hu, X. and Xiao, R., 2025. *Fusion of ECG foundation model embeddings to improve early detection of acute coronary syndromes.* MEDINFO 2025

6. Advances in Language Recognition in Low Resource African Languages: The JHU-MIT Submission for NIST LRE22 (INTERSPEECH 2023) [pdf] (Oral)

7. Shao, Y., Villalba, J., Joshi, **S., Kataria**, S., Dehak, N. and Khudanpur, S., *Chunking Defense for Adversarial Attacks on ASR.* INTERSPEECH 2022 [pdf] (Special session)

8. Joshi, S., **Kataria, S.**, Shao, Y., Zelasko, P., Villalba, J., Khudanpur, S. and Dehak, N., *Defense against Adversarial Attacks on Hybrid Speech Recognition using Joint Adversarial Fine-tuning with Denoiser.* INTERSPEECH 2022 [pdf]

9. Villalba, J., Borgstrom, B.J., **Kataria, S.**, Rybicka, M., Castillo, C.D., Cho, J., Garcıa-Perera, L.P., Torres-Carrasquillo, P.A. and Dehak, N., *Advances in Cross-Lingual and Cross-Source Audio-Visual Speaker Recognition: The JHU-MIT System for NIST SRE21.* Odyssey 2022. [pdf]

10. Villalba, J., Borgstrom, B.J., **Kataria, S.**, Cho, J., Torres-Carrasquillo, P.A. and Dehak, N., *Advances in Speaker Recognition for Multilingual Conversational Telephone Speech: The JHU-MIT System for NIST SRE20 CTS Challenge.* Odyssey workshop 2022 [pdf]

11. Nidadavolu, P.S., **Kataria, S.**, Villalba, J., Garcia-Perera, P. and Dehak, N., *Unsupervised feature enhancement for speaker verification.* ICASSP 2020 [pdf]

12. Villalba, J., Garcia-Romero, D., Chen, N., Sell, G., Borgstrom, J., McCree, A., Garcia-Perera, L.P., **Kataria, S.**, Nidadavolu, P.S., Torres-Carrasquillo, P.A. and Dehak, N., *Advances in speaker recognition for telephone and audio-visual data: the JHU-MIT submission for NIST SRE19.* Odyssey 2020 [pdf]

13. Nidadavolu, P.S., **Kataria, S.**, Villalba, J. and Dehak, N., *Low-resource domain adaptation for speaker recognition using cycle-gans.* ASRU 2019 [pdf]

14. García, P., Villalba, J., Bredin, H., Du, J., Castan, D., Cristia, A., Bullock, L., Guo, L., Okabe, K., Nidadavolu, P.S. and **Kataria, S.**, *Speaker detection in the wild: Lessons learned from JSALT 2019.* Odyssey 2020 [pdf]

15. Gaultier, C., **Kataria, S.** and Deleforge, A., *VAST: The virtual acoustic space traveler dataset.* In International Conference on Latent Variable Analysis and Signal Separation (ICA-LVA 2015) [pdf] (Released dataset)

16. Parthasarathy, A., **Kataria, S.**, Kumar, L. and Hegde, R.M., *Representation and modeling of spherical harmonics manifold for source localization.* ICASSP 2015 [pdf] (Oral)

ABSTRACTS

1. Brown, S.R., Xie, Y., Grunwell, J., Hu, X., Yang, C., **Kataria, S.**, Bold, D., and Lu, J., *Early Prediction of Pediatric In-Hospital Cardiac Arrest Using a Tabular Foundation Model* (Critical Care Congress by Society of Critical Care Medicine (SCCM'25)) [link]

2. Brown, S.R., Xie, Y., Grunwell, J., Hu, X., Yang, C., **Kataria, S.**, Bold, D., and Lu, J., *External Validation of a Model to Predict Cardiac Arrest in the Pediatric Intensive Care Unit*, (Pediatric Research for the Future of Children conference 2026)

3. **Kataria, S.**, Ermis, A., Panchumarthi, L.Y., Wang, M. and Hu, X., 2025. *Vision4PPG: Emergent PPG Analysis Capability of Vision Foundation Models for Vital Signs like Blood Pressure.* (IEEE BHI'25) [pdf]

4. **Kataria, S.**, Lu, J., Xiao, R., Ruchti, T., Clark, M., Bergmann, J. and Hu, X., 2025. Cardiac arrest classification using large PPG foundation model representations. (International Conference on Complex Acute Illness (ICCAI'24)) [pdf]

UNDER REVIEW MANUSCRIPTS

1. **Kataria, S.**, Wu, Y., Chen, Z., Kwak, H.G., Xu, Y., Panchumarthi, L.Y., Xiao, R., Lu, J., Ermis, A., Zhao, A. and Yan, R., *Generalist vs Specialist Time Series Foundation Models: Investigating Potential Emergent Behaviors in Assessing Human Health Using PPG Signals.* In submission to *Nature Communications.* [pdf]
2. **Kataria, S.**, Fattahi, D., Wang, M., Xiao, R., Clark, M., Ruchti, T., Mai, M. and Hu, X., *Wav2Arrest 2.0: Long-Horizon Cardiac Arrest Prediction with Time-to-Event Modeling, Identity-Invariance, and Pseudo-Lab Alignment.* In submission to *Biomedical Signal Processing and Control.* [pdf]

PARTICIPATION IN CHALLENGES

(in collaboration with MIT Lincoln Laboratory, Panasonic, etc.)
NIST Language Recognition Evaluation (LRE): '22, '17 • NIST Speaker Recognition Evaluation (SRE): '21, '20, '19, '18

TALKS

"Robust Speaker Verification using perceptual and adversarial speech enhancement", Brookhaven National Lab, NY, May'23
"Conditioning bandwidth extension GAN models with self-supervision", CLSP seminar, JHU, Feb'23
"CycleGAN variants for Speech Domain Adaptation and Bandwidth Expansion", CLSP seminar, JHU, Nov'21
"Single-Channel Speech Enhancement with Deep Feature Losses for Speaker Verification", CLSP seminar, JHU, Feb'20

TEACHING/COURSE ASSISTANT

Machine Learning for Medical Applications (2021, EN.520.639, JHU): Developed new assignments (incl. Bio BERT)
Computational Modeling for Electrical and Computer Engineering (2020, EN.520.123, JHU)
Machine Learning for Signal Processing (2020, EN.520.612, JHU)
Modeling and Representation Techniques for Images (2017, EE698K, IIT Kanpur)
Image processing (2016, EE604A, IIT Kanpur)

REVIEWER

Engineering Research Express ('25), JMIR Medical Informatics journal ('25), ICASSP ('26, '25, '23), Engineering Applications of Artificial Intelligence (EAAI'25), Sensors journal ('25), TMLR (multiple), Multimedia Systems (Journal), BHI'25, The Journal of Supercomputing, TDSC (IEEE Transactions on Dependable and Secure Computing), TASLP (IEEE/ACM Transactions on Audio, Speech, and Language Processing) (multiple), INTERSPEECH ('25-'23), PMEA journal (multiple), IJCNN'25, BMVC'24, ICLR ('23, '21), CISS ('23), ACML ('22), Speech Communication (multiple), NeurIPS'20

REFERENCES

Prof. Xiao Hu (Emory), Prof. Ran Xiao (Emory), Prof. Jesus Villalba (JHU)