Swapnil Bhosale

swapb94.github.io/ | Guildford, UK
+447833862119
s.bhosale@surrey.ac.uk
vbhosale.bhosale@gmail.com
linked.in/swapnil-bhosale | G Scholar

Research Interests:
Audio-visual correspondence learning,
Novel view synthesis, LLMs

Education

University of Surrey - People Centred AI Institute — United Kingdom
PhD - Vision, Speech, Signal Processing — *2022 – Present*
Focus: Audio-visual correspondence learning.

Walchand College of Engineering — India
Bachelor of Technology in Computer Science and Engineering — *2015 – 2019*
Thesis: End-to-End spoken language understanding

Research Experience

Research Scientist Intern — Sep. 2025 – Present
Mitsubishi Electric Research Laboratories (MERL) — *Cambridge, US*
- Spatial room impulse response modeling.
- 3D Audio-visual correspondence learning.

Research Scientist Intern — Mar. 2025 – Aug. 2025
Meta Reality Labs Research — *Cambridge, UK*
- Real-time sound recognition for smart glasses.
- Encoding acoustic awareness for contextual AI.

Research fellow - Search and Sciences team — Sep. 2023 – Dec. 2023
Ebay Research — *Guildford, UK (Remote)*
- Enhancing product retrieval for e-commerce by optimizing the user intent in semantic product search.
- Curated challenging evaluation sets for implementing user-intent centrality, resulting in significant product ranking efficiency improvements.

Researcher - Speech and NLP Team — Aug. 2019 – Sep. 2022
TCS Research and Innovation — *Mumbai, India*
- Robust audio event detection systems for non-stationary distributions.
- Fusing multimodal cues for emotion recognition.
- Pathological speech processing, focus: Automated intelligibility assessment.

Research Intern - Speech and NLP Team — Jan. 2019 – May 2019
TCS Research and Innovation — *Mumbai, India*
- End-to-End spoken language understanding (SLU) for low-resource scenarios.
- SLU domain adaptation for disordered speech.

Publications - Conferences

1. **Bhosale, Swapnil**, Haosen Yang, Diptesh Kanojia, Jiankang Deng, and Xiatian Zhu. Unsupervised audio-visual segmentation with modality alignment. In *Association for the Advancement of Artificial Intelligence (AAAI) 2025* (**link**)
2. **Bhosale, Swapnil**, Haosen Yang, Diptesh Kanojia, Jiankang Deng, and Xiatian Zhu. AV-GS: Learning material and geometry aware priors for novel view acoustic synthesis. *Advances in Neural Information Processing Systems (NeurIPS)*, 2024 (**link**)
3. **Bhosale, Swapnil***, Sauradip Nag*, Diptesh Kanojia, Jiankang Deng, and Xiatian Zhu. DiffSED: Sound Event Detection with Denoising Diffusion. In *Association for the Advancement of Artificial Intelligence (AAAI) 2024 - Oral* (**link**)
4. Hadeel Saadany, **Bhosale, Swapnil**, Samarth Agrawal, Constantin Orasan, Zhe Wu, and Diptesh Kanojia. Centrality-aware product retrieval and ranking. In *Conference on Empirical Methods in Natural Language Processing (EMNLP): Industry Track*, 2024 (**link**)
5. Artem Sokolov, **Bhosale, Swapnil**, and Xiatian Zhu. 3D Audio-visual Segmentation. *Advances in Neural Information Processing Systems (NeurIPS) Workshop: Audio Imagination*, 2024 (**link**)
6. **Bhosale, Swapnil**, Haosen Yang, Diptesh Kanojia, and Xiatian Zhu. Leveraging foundation models for unsupervised audio-visual segmentation. *IEEE/CVF International Conference on Computer Vision (ICCV) Workshop: AV4D*, 2023 (**link**)
7. **Bhosale, Swapnil**, Rupayan Chakraborty, and Sunil Kumar Kopparapu. Text-to-audio grounding based novel metric for evaluating audio caption similarity. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2023* (**link**)
8. **Bhosale, Swapnil***, Upasana Tiwari*, Rupayan Chakraborty, and Sunil Kumar Kopparapu. Contrastive Learning of Cough Descriptors for Automatic COVID-19 Preliminary Diagnosis. In *Special Session DiCOVA at Interspeech 2021* (**link**)
9. **Bhosale, Swapnil***, Upasana Tiwari*, Rupayan Chakraborty, and Sunil Kumar Kopparapu. Deep Lung Auscultation Using Acoustic Biomarkers For Abnormal Respiratory Sound Event Detection. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2021* (**link**)
10. **Bhosale, Swapnil**, Rupayan Chakraborty, and Sunil Kumar Kopparapu. Deep Encoded Linguistic and Acoustic Cues for Attention Based End to End Speech Emotion Recognition. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2020* (**link**)
11. **Bhosale, Swapnil**, Imran Sheikh, Sri Harsha Dumpala, and Sunil Kumar Kopparapu. End-to-End Spoken Language Understanding: Bootstrapping in Low Resource Scenarios. In *Interspeech 2019* (**link**)
12. Ayush Tripathi, **Bhosale, Swapnil**, and Sunil Kumar Kopparapu. Improved Speaker Independent Dysarthria Intelligibility Classification Using Deepspeech Posteriors. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2020* (**link**)
13. Ayush Tripathi, **Bhosale, Swapnil**, and Sunil Kumar Kopparapu. A Novel Approach for Intelligibility Assessment in Dysarthric Subjects. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2020* (**link**)
14. Ayush Tripathi, **Bhosale, Swapnil**, and Sunil Kumar Kopparapu. Automatic Speech Intelligibility Assessment in Dysarthric Subjects (Demo). In *The Fourteenth International Conference on Digital Society*. IARIA, 2020 (**link**)

Journal Papers

1. **Bhosale, Swapnil**, Rupayan Chakraborty, and Sunil Kumar Kopparapu. Calibration free Meta learning based approach for Subject Independent EEG Emotion Recognition. In *Biomedical Signal Processing and Control 2022.* (**link**)

2. Ayush Tripathi, **Bhosale, Swapnil**, and Sunil Kumar Kopparapu. Automatic Speaker Independent Dysarthric Speech Intelligibility Assessment System. *Computer Speech & Language*, 69, 2021 **(link)**

Patents

1. **Bhosale, Swapnil**, Rupayan Chakraborty, Sanat Sarangi, Sanket Kailas Junagade, and Srinivasu Pappula. Methods and systems for generating optimized planting schedule of crop to overcome storage capabilities, May 30 2024. US Patent App. 18/376,648
2. **Bhosale, Swapnil**, Rupayan Chakraborty, and Sunil Kumar Kopparapu. Methods and Systems For Building A SemiSupervised Few-Shot Model. Indian Patent No. 553611, Granted 30 October 2024
3. Tripathi, Ayush and **Bhosale, Swapnil** and Kopparapu, Sunil Kumar. Methods and Systems For Assessment of Speech Intelligibility in Dysarthric Subjects. Indian Patent No. 549763, Granted 06 September 2024.

Technical Skills

Languages: Python, R, HTML/CSS
Deep learning Toolkits: PyTorch, Keras, Tensorflow
Frameworks: Flask
Developer Tools: Git, VSCode

Teaching Experience

EEEM066: Fundamentals of Machine Learning – Lab Fall 2023
- Linear Algebra, Neural Networks, Machine Learning System Design.

EEEM071: Advanced Computer Vision and Deep Learning – Lab Spring 2023 | Spring 2024
- Image representation, CNN Interpretability, Transformers, Domain adaptation.

EEE1032: Mathematis II - Engineering Mathematics Spring 2023
- Signal Theory, Engineering Mechanics, Ordinary Differential Equations

Languages

English: Fluent (IELTS: Band-8)
Marathi: Fluent (Native)
Hindi: Fluent