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

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: 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.

Machine Learning Intern — Mar. 2018 – Sep. 2018
Chainrule.ai — *Remote*
- Neural approaches for chest Xray cancer detection and segmentation.
- Brain tumor segmentation.

Publications - Conferences

1. **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**)
2. **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**)
3. **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**)
4. **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**)
5. **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**)

6. **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**)
7. **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**)
8. **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**)
9. 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**)
10. 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**)
11. 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. Tripathi, Ayush and **Bhosale, Swapnil** and Kopparapu, Sunil Kumar. Methods and Systems For Assessment of Speech Intelligibility in Dysarthric Subjects. Indian Provisional Patent Ser. No. 202021008649, Granted 05 September 2024
3. **Bhosale, Swapnil**, Rupayan Chakraborty, and Sunil Kumar Kopparapu. Methods and Systems For Building A SemiSupervised Few-Shot Model. Indian Provisional Patent Ser. No. 202021034689, filed 12 August 2020

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

Achievements

Predict-X, Mindspark-College of Engineering, Pune - Winner Sep. 2017
- Category : *Time Series Prediction, Natural Language Processing (NLP)*
- Involved two problem statements, Uni-variate Time Series prediction on stock prices and developing a rating system for products by sentiment analysis on Amazon product reviews.

Smart India Hackathon 2018 - Finalist Apr. 2018
- Category : *Recommendation system, Clustering, Churn Prediction*
- Developed an Online Incubator platform for Skill Development and Entrepreneurship Ministry of India. Built recommendation systems based on location and job preferences for registered members.

E-Yantra, IIT Bombay - Semi Finalist 2017 - 2018
- Category : *Path planning, swarm robotics, Arduino programming*
- Built a system of fully autonomous weeder bots using ATMega256 controller, in an attempt to implement the concepts of swarm robotics for farming.

AWS Deep Learning Hackathon, IIT Madras - Placed 6th Jan. 2018
- Category : *Deep Learning, Machine Learning, Feature Engineering*
- Task 1: Developed a driver distraction detection system with Geo-tagging deployed using AWS lambda endpoints.
- Task 2: Predict the maintenance date for Rolls Royce water pumps using past history and process logs.

National Robotics Championship 2016, IIT Bombay - Zonal Winner Mar. 2016
- Category : *Arduino programming, Micro controllers*
- Developed Arduino based autonomous line follower bots.

Projects

Web Template Image to code Generation | *Python, Flask, Keras*
* Trained an End-to-End model to jointly learn the spatial characteristics from a snapshot (image) of a webpage and sequential information from its corresponding HTML code template.
* Encoded vectors for image and text (code) are generated using two separate convolutional and RNN encoders respectively, and later fused using an attention mechanism

Face Sketch to Photo-realistic Images using GANs | *Python, Flask, Keras*
* Trained a model based on Deep convolutional conditional Generative Adversarial Networks, to generate realistic images from hand-drawn face sketches.
* Later extended to generate enhanced images from poorly lit images, or blurry images.
* Incorporated separate gender prediction and age estimation models over the generated image, -particularly, important from the forensic identification perspective.

Driver Distraction Detection System | *Python, TKinter, Keras*
* Built a system to detect a distracted driver and alert him/her through a computer generated voice.
* The system used a model trained using CNNs to classify the driver's actions into a set nine predefined distractions. The model took multiple frames from a streaming video feed captured from a dashboard camera.

Languages

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