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:
Multimodal learning, Meta learning,
Audio and Speech Processing, NLP

Education

Surrey AI Institute, University of Surrey — UK
PhD - Vision, Speech, Signal Processing — *Sep. 2022 – Present*

Walchand College of Engineering — MH-India
Bachelor of Technology in Computer Science and Engineering — *May 2015 – May 2019*

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. *(ongoing)*

- Proposed a semi-supervised few-shot learning paradigm, for robustness towards unseen classes by mining episodic triplets as opposed to conventional triplet loss.
- Adapting continual learning architectures to classify novel classes, while simultaneously retaining classification performance towards base classes.
- Substantial improvements observed over existing supervised models as well as existing few-shot approaches. *Outcome:* TCS-FiSound, A system for detecting audio patterns in the wild; [(pre-print)1].

Fusing multimodal cues for Emotion recognition

- Improved speech emotion performance by combining acoustic features with linguistic cues obtained from intermediate layers of pre-trained Speech-to-Text models.[3]
- Proposed a calibration free EEG signal based emotion recognition method by consolidating existing meta-learning based models with a multi-head attention mechanism to learn subject invariant embeddings, capable of incorporating spatio-temporal relationships across adjacent EEG electrodes. [1]

Pathological Speech Processing: Automated Intelligibility Assessment

- Developed an intelligibility assessment system for dysarthric subjects by combining acoustic features with patterns within erroneous ASR transcriptions. Designed a cost minimization approach to quantify minimum number of words required for reliable intelligibility estimation using acoustic features, and Visible Speech representation.[4, 5]

Research Intern - Speech and NLP Team — Jan. 2019 – May 2019
TCS Research and Innovation — *Mumbai, India*

Spoken Language Understanding for low-resource scenarios.

- Explored transfer learning for low resource spoken language understanding without Speech-to-Text conversion, by localizing speaker invariant features learned by intermediate layers of pre-trained ASR models.
- Substantial improvements achieved over conventional MFCC-based features even when evaluated in language mismatched test setups.[6]

Machine Learning Intern — Mar. 2018 – Sep. 2018
Chainrule.ai (now, Datanudge.ai) — *Remote*

Automated Classification from chest Radiographs using Deep Learning.

- Developed classification and segmentation models using Deep Learning for medical data(X-Rays, CT scans, MRI scans).
- Majorly contributed in,
  1. Chest Xray cancer detection and segmentation [NIH Dataset]
  2. Pancreatic Cancer Survival Rate Prediction [MICCAI 2018-PDAC]
  3. Brain Tumor Segmentation Challenge [BraTS 2018]

Summer Research Intern Jun. 2018 – Jul. 2018
TCS Research and Innovation *Mumbai, India*

Speech Command Recognition using End-to-End Deep models.

- Carried out experiments and analysis on Deep Learning models for classifying spoken commands and spoken utterances (on lines of spoken language understanding without ASR).
- Experimented with various architectures - Convolutional Neural Networks, LSTMs, CNN-LSTMS, Hierarchical Networks on speech datasets.
- Mainly focused on deriving meaningful intuitions by visualizing the layer weights, and using manifold learning algorithms(PCA,T-SNE) for dimensionality reduction.

Publications

1. **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**)
2. Upasana Tiwari*, **Bhosale, Swapnil***, 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**)
3. **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*, pages 7189–7193 (**link**)
4. 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*, pages 6114–6118 (**link**)
5. 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*, pages 6779–6783 (**link**)
6. **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*, pages 1188–1192 (**link**)
7. 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.*
2. Ayush Tripathi, **Bhosale, Swapnil**, and Sunil Kumar Kopparapu. Automatic Speaker Independent Dysarthric Speech Intelligibility Assessment System. *Computer Speech & Language*, 69:101213, 2021 (**link**)

Pre-prints

1. **Bhosale, Swapnil**, Rupayan Chakraborty, and Sunil Kumar Kopparapu. Semi Supervised Learning For Few-shot Audio Classification By Episodic Triplet Mining. *arXiv preprint arXiv:2102.08074*, 2021 (**link**)

Patents - Filed

1. **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
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, filed 28 February 2020

Technical Skills

Languages: Python, R, Java, C/C++, JavaScript, HTML/CSS
Deep learning Toolkits: PyTorch, Keras, Tensorflow
Frameworks: Flask, Django
Developer Tools: Git, PyCharm

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 (Dt: Oct 24, 2020))
Marathi: Fluent (Native)
Hindi: Fluent