

## Sarath Sivaprasad

+91 9497457367 | [sarath.s@research.iiit.ac.in](mailto:sarath.s@research.iiit.ac.in) | [Website](#)

### EDUCATION

---

**International Institute of Information Technology (IIIT), Hyderabad** July 2019 – March 2022  
*MSc-Research in Computer Science*

**Amrita School of Engineering (ASE), Amrita Viswa Vidyapeetham** July 2011 – March 2015  
*BTech in Electronics and Communication Engineering*

### PUBLICATIONS

---

- “Adversarial Robustness of Mel based speaker recognition systems”, Submitted at **ICASSP 2022**
- “The Curious Case of Convex Neural Networks”, **ECML-PKDD, 2021**
- “Reappraising Domain Generalization in Neural Networks”, **arXiv preprint 2021**
- “Emotional Prosody Control for Speech Generation”, **Interspeech, 2021**
- “Color Me Good: Rendering Logos in the Coloring Style of Movie Posters”, **Workshop at CVPR 2021**
- “Partners in Crime: Utilizing A-V Relationship for Continuous Prediction of Valence in Movies”, **AffCon, AAAI 2019**
- “Multimodal Continuous Prediction of Emotions in Movies using Long Short-Term Memory Networks”, **ICMR 2018**
- “Multimodal Approach to Predicting Media Memorability”, **MediaEval 2018**

### TECHNICAL SKILLS

---

Machine Learning, Mathematical Modeling, Technical Writing, Multimodal Analysis: Vision, Text & Audio Technologies: Python (Tensorflow, Pytorch), Matlab, C, C++, Verilog, MySQL

### RESEARCHER @ CENTER FOR VISUAL INFORMATION TECHNOLOGY (CVIT)

---

**Research towards master’s thesis** July 2019- present  
*IIIT-Hyderabad* *Hyderabad, India*

#### **Input Output Convex Neural Networks (IOC-NNs) for improved generalization performance**

- Proposed a method to constrain output of NNs as convex function of its inputs, whilst retaining adequate capacity
- IOC-NNs self regularize and give significant reduction in empirical error, reducing the risk of over-fitting
- Demonstrated IOC-NNs’ robustness to label noise and improvement in confidence calibration
- IOC constraint improves test performance for MLPs and gives competitive performance on CNN architectures

#### **Domain Generalization (DG) without bells and whistles: Revisiting challenges in DG**

- Proposed a training protocol which gives SOTA performance on four standard DG benchmarks
- A simple ERM with a chosen backbone (chosen from ablation study) outperforms methods tailored for DG
- Proposed an alternate evaluation framework (ClasswiseDG) where, a domain from each class is kept out for testing
- Despite being exposed to all the domains during training, NNs find it challenging to perform in this framework

#### **Disentangling spurious variances from ‘meaningful’ ones**

- Encoders learn same representation for image and its background masked pair, disregarding background variances
- Working on improving the balance of texture and shape bias of neural network using modified self challenging
- Compare against humans the average number of images NN requires to spot an object from multi object images

#### **Text to Speech (TTS) conversion conditioned on emotions**

- Built on FastSpeech2 framework to develop a TTS system that can render speech in a given emotion
- Variance prediction (pitch, energy & duration) is modelled to hinge on emotion from a human interpretable space
- Used Speaker discriminative embedding to attend on phonemes to scale TTS to multiple speakers

#### **Behaviour aware conversational BoT: variance scaling from emotion cues**

- Modelled response of an agent in a conversation with ‘behaviour’ as the apparent and ‘emotion’ as a latent state
- Improved our TTS model using differential scaling of variance in the direction of predicted emotion
- Showed that a change in initial state of an agent can cause large variation in trail of emotions in a dialogue

## RESEARCHER @ TATA RESEARCH DEVELOPMENT AND DESIGN CENTER (TRDDC)

---

**Multimedia, Machine Learning and Behavioural Science**

May 2017 - present

*TRDDC, Pune*

*Pune, India*

### **Multimodal intensity profiling of media**

- Used hand crafted audio and video features to predict the emotional content (Arousal - Valence) of movies
- Used LSTM based MoEs to capture the temporal context. Reported results that beat baseline by 20%
- Modelled the covariance of arousal with valence as a regularization, reported further improvement of 20%

### **Advertisement understanding for better Ad placement in entertainment media content**

- Designed an ad placement platform to minimize user distraction, using affective annotations of media content
- Understand ads by predicting the persuasion and reasoning presented through symbolism, improved SOTA method
- Ad Image is style transferred to blend into the least salient corner of scene that matches with emotion of the ad

### **Automatic prediction of multimedia content memorability**

- Used image, video and text based features to predict memorability score for short clips. Dataset: MediaEval
- Modelled memorability as a function of high level modalities like saliency, aesthetics, activity and text summary
- Found text summary as the most salient feature. Proposed a metric derived from Tf-Idf to capture text relevance

## RESEARCHER & LEARNING ENABLER @ TCS ILP INNOVATION LAB

---

**Lead, AI development team and ML Tutor**

Aug 2016 - Sept 2015

*TCS, ILP center Trivandrum*

*Trivandrum, India*

### **Comprehensive compilation and content indexing in product specification documents**

- Developed a system to organize and index information from volumes of product specification documents
- Extracted information about figures from text, and developed a CNN based system to read data from graphs
- Labelled images with relevant text using CNNs and created a database which can be queried with keywords

### **Language assessment automation**

- Developed a tool to grade the writing and speaking proficiency, trained on previous expert evaluations
- Score computed using hand crafted features based on spelling, grammar, speech rate, recurring words, pausing etc.
- Used LSTM based models to grade coherence and clarity in writing and mother tongue influence in speech

### **Authentication System and Method for forensic analysis (Patent Granted: US 9916511)**

- Utilized CNN activations at selected layers of the input identification image as a signature matrix
- Compared the matrix against the database of pre-computed activations to get samples for fine search
- Used a variant of triplet loss to ensure uniqueness of activations generated at selected layers

## SELECTED AWARDS AND ACHIEVEMENTS

---

- Stood first in 'Annual Implementation Hackathon 2016' of manufacturing unit CTO lab, TCS
- Awarded 'Employee of the quarter', 'Spot award', 'Client appreciation' and 'Exceptional contribution award' at TCS
- The IEEE Student Branch ASE won the best IEEE SB award during my tenure as the office bearer
- Stood first in the national robotics competition at IIST Trivandrum and NIT Calicut

## SOCIALLY IMPACTFUL PROJECTS

---

- At **Humanitarian Technology Lab**, developed low cost solution for healthcare gadgets to improve it's accessibility. Prototyped Pulse Oximeter, Ultrasound Scanner and gesture recognition module for Automatic Wheelchair.
- Developed an **E-Governance mobile platform (M-Kerala)** that allows citizens to submit queries and a web portal for administrators to update status. Received medal of appreciation from Chief Secretary of State, Kerala.
- Developed a **VR Game to help rehabilitation of autism affected children**. Modelled VR game variables based on emotive sensor outputs. Received positive feedback from doctors about the interest and engagement of children.
- Developed and deployed a **Big-data visualization platform for the State Police force** of Kerala to detect anomaly in social media feeds. Our team stood first in Startup mission Kerala hackathon.