Ashita Batra

Ph.D. Scholar • Department of Computer Science & Engineering • IIT Guwahati • Speech Lab | SPIN Lab

(+91) 9675736879 | b.ashita@iitg.ac.in | ashitabatra23

Education

2020-Pursuing	Ph.D (Computer Science & Engg.)	Indian Institute of Technology, Guwahati	**7.5/10.0**
2018-2020	M.Tech (Computer Science & Engg.)	National Institute of Technology, Hamirpur	**8.76/10.0**
2013-2017	B.Tech (Computer Science & Engg.)	Uttarakhand Technical University	**79.10%**

Research Experience

Automatic assessment and detection of stuttering in spoken conversations | *Guided by - Prof. Pradip K. Das & Dr. Neeraj Kumar Sharma*

- Automatic assessment and detection of stuttering in spoken conversations
- Quantify the acoustic differences between fluent and disfluent speech
- Quantifying the cognitive load experienced by individuals who stutter while having conversations
- Data collection for regional languages like Hindi, Bengali, Assamese.
- Studying the causes of neurological disorder which affects speech from engineering point of view
- Make a multi-modal and multi-lingual architecture which can differentiate between various stuttering types

Multi-Focus Image fusion (M.Tech Thesis) | *Guided by - Dr. Prakash Choudhary*

- Exploited Wavelet Transform for the multi-focused input images which takes low and high-level coefficients as input images
- Implemented CNN architecture to extract features low-level & high-level coefficients to obtain fused image
- Implemented hybrid technique using wavelet transform for pre-processing of images and convolution neural network based architecture for end-to-end image fusion of multi-focus images

Publications

- **An Indian Stuttered Speech Dataset to Detect Stuttering Events in Read and Spontaneous speech.** *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2025)*
- **Does Data Balancing Impact Stutter Detection and Classification?** *(21st International Conference on Distributed Computing and Intelligent Technology (ICDCIT 2025))*
- **DDoS: Detecting different types of stuttered speech of multiple speakers exploiting various feature extraction techniques.** *(6th International Conference on Machine Learning, Image Processing, Network Security and Data Sciences (MIND) 2024)*
- **Skin Patch based Age Detection using Machine Learning and Transfer Learning approach.** *(Communicated in Multimeda Tools and Applications (MTAP))*
- **Comparative Analysis of Classifiers using Wav2Vec2.0 Layer Embedding for Imbalanced Stuttering Datasets.** International Conference on Electronics, Communication and Signal Processing (2024).
- **Machine Learning Models Based Stuttering Classification** Innovations in Computational Intelligence and Computer Vision (2024).
- **A review on automatic assessment and detection of pathological speech** In: Swain, B.P., Dixit, U.S. (eds) Recent Advances in Electrical and Electronic Engineering. ICSTE 2023. Lecture Notes in Electrical Engineering, vol 1071. Springer, Singapore. $https://doi.org/10.1007/978-981-99-4713-3\_40$
- **VTnet+Handcrafted based approach for food cuisines classification.** Multimeda Tools and Appl (2023). $https://doi.org/10.1007/s11042-023-15800-4$.
- **Food Classification of Indian Cuisines Using Handcrafted Features and Vision Transformer Network**. SSRN. $http://dx.doi.org/10.2139/ssrn.4014907$.
- **A Deep Learning Hybrid CNN Framework Approach for Vegetation Cover Mapping Using Deep Features**. Proceedings of 13^{th} International Conference on Signal Image Technology & Internet-Based Systems (SITIS), 2017. $https://doi.org/10.1109/SITIS.2017.41$

Key Projects

Medical Prescription Generator using speech signals | ***Based of HMM***

- Developed a user-specific speech based prescription generator that recognizes the symptoms spoken by the user and generates prescription accordingly.
- Recognition was done with the help of LPC coefficients and Hidden Markov Model(HMM) was developed from scratch in C.
- Users could also add new words live in the system.

Speaking calculator using analysis of speech signals | ***Based on HMM***

- Speaking Calculator was made using the HMM(Hidden Markov model) model.
- The training was done on two different voices. The voice on which the dataset was made, went through various stages of data normalization, denoising etc.
- The testing was done in real-time wherein the numbers with associated operation has to be performed was done and the output was also provided in real-time with an accuracy of 70%.

Sentiment Analysis | ***Based on NLP***

- Analyzed thousands of twitter tweets to predict people's sentiment
- Data cleaning done by removing punctuations,stopwords and then performing tokenization
- After creating a pipeline for data-cleaning process.Trained the model to make predictions and then classified using Naive Bayes classifier

Traffic sign classification for self-driving cars | ***Based on Deep Learning using Keras***

- Performed image normalization and convert from color-scaled to gray-scaled images
- Built a Convolutional Neural Network using Keras with Tensorflow 2.0 as a backend
- Compile and fit Deep Learning model to training data
- Assess the performance of trained CNN and ensure its generalization using various KPIs.
- Improve network performance using regularization techniques such as dropout

A Deep Learning hybrid CNN framework approach for age estimation using skin patches |

Based on Deep Learning

- Did age estimation using the collected dataset of approximately 2k images belonging to different classes.
- A hybrid architecture of CNN for feature extraction and SVM for classification has been used.
- Comparartive analysis has been done on various architectures of CNN like MobileNet,AlexNet,VGG19 to check which model performs better using the concept of transfer learning.
- Accuracy metric is used for evaluation of the models.It was observed that MobileNet performs best amongst all.

Skills

Tools • MYSQL • Oracle 10g • NoSQL • LaTeX

Subjects • Operating system • Algorithms • Programming & Data Structures • Discrete Mathematics • Speech processing • Linear Algebra • Database Management System • Statistics

Libraries • NumPy • Pandas • Keras • Tensorflow • PyTorch • HuggingFaces

Frameworks • RaWNet • ResNet • VGGNet • Wav2Vec2.0 • BERT • TDNN • Whisper

Toolkit • SpeechBrain • Praat • Audacity • Kaldi