Saurabhchand Bhati

3925 Beech Avenue
Baltimore, MD, USA - 21211

Phone: +1(443)630-2424
email: *sbhati1@jhu.edu*

EDUCATION

Johns Hopkins University, USA — 2018-ongoing
PhD candidate, Electrical and Computer Engineering
Research Interests: Unsupervised learning, Speech processing, Deep learning

Indian Institute of Technology Hyderabad, India — 2012-2017
B.Tech.(Hons.)-M.Tech Dual Degree, Department of Electrical Engineering

WORK EXPERIENCE

Amazon Alexa: Applied Scientist Intern — May'21-Aug'21
- Large scale self-supervised learning for Bilingual ASR

Indian Institute of Technology Hyderabad: Senior Research Scientist — May'17-Aug'18
- Unsupervised acoustic unit discovery for language identification

RELEVANT PUBLICATIONS [SCHOLAR]

Journal Papers

- **Bhati S.**, Villalba, J., Żelasko P., Moro-Velazquez L. and Dehak, N., "*Unsupervised Speech Segmentation and Variable Rate Representation Learning using Segmental Contrastive Predictive Coding*," IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2022. [pdf]
- Żelasko P., Feng S., Velázquez L.M., Abavisani A., **Bhati S.**, Scharenborg O., Hasegawa-Johnson M. and Dehak N., "*Discovering phonetic inventories with crosslingual automatic speech recognition*," Computer Speech & Language, 2022. [pdf]
- **Bhati S.**, Nayak S. and Murty, K.S.R., "*Unsupervised Speech Signal to Symbol Transformation for Language Identification*," Circuits, Systems, and Signal Processing, 2020. [pdf]
- **Bhati S.**, Villalba, J., Moro-Velazquez L., Żelasko P. and Dehak N., "*Regularizing Contrastive Predictive Coding for speech applications*," in submission to IEEE/ACM TASLP, 2023. [pdf]

Conference Papers

- **Bhati S.**, Villalba, J., Moro-Velazquez, L. and Dehak, N. "*Segmental SpeechCLIP: Utilizing Pretrained Image-text Models for Audio-Visual Learning.*" in submission, 2023 [pdf]
- Khare, A., Wu, M., **Bhati S.**, Droppo, J. and Maas, R., "*Guided contrastive self-supervised pretraining for automatic speech recognition.*" in IEEE Spoken Language Technology, 2022 [pdf]
- **Bhati S.**, Villalba, J., Żelasko, P., Moro-Velazquez, L. and Dehak, N. "*Segmental Contrastive Predictive Coding for Unsupervised Word Segmentation.*" in Proc. Interspeech - 2021 [pdf]
- **Bhati S.**, Villalba, J., Żelasko, P., and Dehak, N. "*Self-expressing autoencoders for unsupervised spoken term discovery.*" in Proc. Interspeech - 2020 [pdf]
- **Bhati S.**, Nayak, S., Murty, K.S.R. and Dehak, N, "*Unsupervised Acoustic Segmentation and Clustering Using Siamese Network* " in Proc. Interspeech - 2019 [pdf]
- **Bhati S.**, Velazquez, L.M., Villalba, J. and Dehak, N.. "*LSTM Siamese Network for Parkinson's Disease Detection from Speech.*" In IEEE Global Conference on Signal and Information Processing (GlobalSIP), 2019. [pdf]
- Nayak, S., **Bhati S.** and Murty, K.S.R., "*Zero resource speaking rate estimation from change point detection of syllable-like units*." In IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2019. [pdf]
- **Bhati S.**, Liu, C., Villalba, J., Trmal, J., Khudanpur, S. and Dehak, N., "*Bottom-Up Unsupervised Word Discovery via Acoustic Units*." In IEEE Global Conference on Signal and Information Processing (GlobalSIP), 2019. [pdf]

- Nayak, S., Shashank, D.B., **Bhati S.**, Bramhendra, K. and Murty, K.S.R., "*Instantaneous frequency features for noise robust speech recognition*. In IEEE National Conference on Communications (NCC), 2019. [pdf]
- Nayak, S., Kumar, C.S., Ramesh, G., **Bhati S.** and Murty, K.S.R., "*Virtual phone discovery for speech synthesis without text.*" In IEEE Global Conference on Signal and Information Processing (GlobalSIP), 2019. [pdf]
- **Bhati S.**, Kamper H. and Murty, K.S.R., "*Phoneme Based Embedded Segmental K-Means for Unsupervised Term Discovery*" in International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2018 [pdf]
- Nayak, S., **Bhati S.** and Murty, K.S.R., 2017, November. "*An investigation into instantaneous frequency estimation methods for improved speech recognition features.*" In IEEE Global Conference on Signal and Information Processing (GlobalSIP), 2017. [pdf]
- **Bhati S.**, Nayak S., and Murty, K.S.R., "*Unsupervised Segmentation of Speech Signals Using Kernel-Gram Matrices*," in Proc. NCVPRIPG 2017, Communications in Computer and Information Science, Springer. [pdf]
- **Bhati S.**, Nayak S., and Murty, K.S.R., "*Unsupervised Speech Signal to Symbol Transformation for Zero Resource Speech Application*," in Proc. Interspeech - 2017 [pdf]

RESEARCH PROJECTS

Multi-modal word discovery Ongoing

Advisor: Dr. Najim Dehak, JHU

- Proposed a segmental audio encoder to learn word-like units from speech with the guidance from paired images
- Leverages pretrained models like CLIP for image representations
- Better retrieval performance than models like cascadedCLIP (pytorch)

Regularized Contrastive Predictive Coding Ongoing

Advisor: Dr. Najim Dehak, JHU

- Proposed slow changing constraints for contrastive predictive coding
- Encourages the adjacent feature to be similar to generate piecewise linear features
- Outperforms the baseline CPC in monolingual, crosslingual and multiligual setting (pytorch)

Segmental Contrastive Predictive Coding (SCPC) May'20-Oct'21

Advisor: Dr. Najim Dehak, JHU

- Proposed SCPC: a hierarchical self-supervised model capable of learning mutli-scale information from raw speech
- Neural model capable of doing both phone and word segmentation jointly
- State-of-the-art performance on unsupervised phone and word segmentation (pytorch, kaldi)

Self-expressing Auto-encoders (SEA) for unsupervised feature learning Aug'19-Apr'20

Advisor: Dr. Najim Dehak, JHU

- Proposed SEA to highlight the underlying class information in unsupervised manner
- Outperforms MFCC in unsupervised, partial supervision and supervised scenarios (pytorch, kaldi)

Unsupervised Term Discovery in Speech July '18 - Oct '19

Advisor: Dr. Najim Dehak, JHU

- Developed a system for automatic discovery of word like units, in untranscribed speech
- State-of-the-art performance on Zero Resource 2015 (15 hours) and Zero Resource 2017 (70+ hours) benchmarks (kaldi, keras, Matlab)

Unsupervised Speech Segmentation July '16 - Oct '16

Advisor: Dr. K. Sri Rama Murty, IIT Hyderabad

- Developed a new kernel gram based segmentation method for locating the phoneme boundaries in raw speech signal
- Achieved highest boundary detection rate (6.5 % higher F-score) on TIMIT (Matlab, keras, kaldi)

OTHER PROJECTS

Parkinson's detection from Speech and accelerometer data Aug '18 - Mar '20

Advisor: Dr. Najim Dehak, JHU

- Proposed a Siamese networks based method Parkinson detection from Speech data
- BeatPD challenge for detecting Dyskinesia, Termor from accelerometer data (Pytorch, kaldi)

Instantaneous Frequency based Automatic Speech Recognition July '16 - Feb '17

Advisor: Dr. K. Sri Rama Murty, IIT Hyderabad

- MFCC features are used for speech recognition which don't utilize phase information
- Combined the proposed IFCC features with traditional MFCC to improve ASR performance by 6% relative on TIMIT dataset (Matlab, kaldi, keras)

SKILLS

Programming Languages: Python, bash
Tools/Frameworks: Pytorch, Matlab, keras, kaldi, Tensorflow, Theano, SciKit, HTK

INVITED TALKS

Hidden Markov Models for Speech Recognition
- TEQIP workshop, Rajiv Gandhi Institute of Technology, Kerala, 2017

Self-expressing autoencoders for unsupervised feature learning
- University of Illinois at Urbana-Champaign, 2020

Segmental Contrastive Predictive Coding for Unsupervised Acoustic Segmentation
- ISCA SIGML Seminar Series, 2021

TEACHING EXPERIENCE

Teaching Assistant

- EN.520.612: Machine learning for signal processing, JHU Aug '19 - Dec '19
  Introduces PCA, PPCA, ICA, NMF, GMMs, HMMs, DNNs, RNNs.
- EE7390: Pattern recognition and Machine learning, IIT Hyderabad Aug '15 - Dec '15, Jan '17 - May '17
  Introduces k-means, K-NN, GMMs, HMMs, DNNs, Linear and Fisher discriminators etc.
- EE5370: Introduction to Machine learning, IIT Hyderabad Aug '16 - Dec '16
  Introduces Naive Bayes classifier, Support vector machines, and clustering techniques.

TRAVEL GRANTS

ISCA travel grant for Interspeech 2017
MHRD travel grant for TEQIP workshop
ISCA grant for JSALT workshop

REVIEWING

IEEE/ACM Transactions on Audio, Speech and Language Processing
Journal of Selected Topics in Signal Processing, IEEE
National Conference on Communications