

# Saurabhchand Bhati

3925 Beech Avenue  
Baltimore, MD, USA - 21211

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

EDUCATION	<b>Johns Hopkins University, USA</b> PhD candidate, Electrical and Computer Engineering Research Interests: Unsupervised learning, Speech processing, Deep learning <b>Indian Institute of Technology Hyderabad, India</b> B.Tech.(Hons.)-M.Tech Dual Degree, Department of Electrical Engineering	2018-ongoing    2012-2017
WORK EXPERIENCE	<b>Amazon Alexa:</b> Applied Scientist Intern • Large scale self-supervised learning for Bilingual ASR <b>Indian Institute of Technology Hyderabad:</b> Senior Research Scientist • Unsupervised acoustic unit discovery for language identification	May'21-Aug'21  May'17-Aug'18
RELEVANT PUBLICATIONS <a href="#">[SCHOLAR]</a>	<b>Journal Papers</b> <ul style="list-style-type: none"><li>• <b>Bhati S.</b>, Villalba, J., Želasko P., Moro-Velazquez L. and Dehak, N., “<i>Unsupervised Speech Segmentation and Variable Rate Representation Learning using Segmental Contrastive Predictive Coding</i>,” IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2022. <a href="#">[pdf]</a></li><li>• Želasko P., Feng S., Velázquez L.M., Abavisani A., <b>Bhati S.</b>, Scharenborg O., Hasegawa-Johnson M. and Dehak N., “<i>Discovering phonetic inventories with crosslingual automatic speech recognition</i>,” Computer Speech &amp; Language, 2022. <a href="#">[pdf]</a></li><li>• <b>Bhati S.</b>, Nayak S. and Murty, K.S.R., “<i>Unsupervised Speech Signal to Symbol Transformation for Language Identification</i>,” Circuits, Systems, and Signal Processing, 2020. <a href="#">[pdf]</a></li><li>• <b>Bhati S.</b>, Villalba, J., Moro-Velazquez L., Želasko P. and Dehak N., “<i>Regularizing Contrastive Predictive Coding for speech applications</i>,” in submission to IEEE/ACM TASLP, 2023. <a href="#">[pdf]</a></li></ul> <b>Conference Papers</b> <ul style="list-style-type: none"><li>• Khare, A., Wu, M., <b>Bhati S.</b>, Droppo, J. and Maas, R., “<i>Guided contrastive self-supervised pre-training for automatic speech recognition</i>,” in IEEE Spoken Language Technology, 2022 <a href="#">[pdf]</a></li><li>• <b>Bhati S.</b>, Villalba, J., Želasko, P., Moro-Velazquez, L. and Dehak, N. “<i>Segmental Contrastive Predictive Coding for Unsupervised Word Segmentation</i>,” in Proc. Interspeech - 2021 <a href="#">[pdf]</a></li><li>• <b>Bhati S.</b>, Villalba, J., Želasko, P., and Dehak, N. “<i>Self-expressing autoencoders for unsupervised spoken term discovery</i>,” in Proc. Interspeech - 2020 <a href="#">[pdf]</a></li><li>• <b>Bhati S.</b>, Nayak, S., Murty, K.S.R. and Dehak, N, “<i>Unsupervised Acoustic Segmentation and Clustering Using Siamese Network</i> ” in Proc. Interspeech - 2019 <a href="#">[pdf]</a></li><li>• <b>Bhati S.</b>, Velazquez, L.M., Villalba, J. and Dehak, N.. “<i>LSTM Siamese Network for Parkinson’s Disease Detection from Speech</i>.” In IEEE Global Conference on Signal and Information Processing (GlobalSIP), 2019. <a href="#">[pdf]</a></li><li>• Nayak, S., <b>Bhati S.</b> and Murty, K.S.R., “<i>Zero resource speaking rate estimation from change point detection of syllable-like units</i>.” In IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2019. <a href="#">[pdf]</a></li><li>• <b>Bhati S.</b>, Liu, C., Villalba, J., Trmal, J., Khudanpur, S. and Dehak, N., “<i>Bottom-Up Unsupervised Word Discovery via Acoustic Units</i>.” In IEEE Global Conference on Signal and Information Processing (GlobalSIP), 2019. <a href="#">[pdf]</a></li></ul>	

- 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 multilingual 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 multi-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)

	<b>Unsupervised Speech Segmentation</b> <span style="float: right;">July '16 - Oct '16</span> <i>Advisor: Dr. K. Sri Rama Murty, IIT Hyderabad</i> <ul style="list-style-type: none"> <li>Developed a new kernel gram based segmentation method for locating the phoneme boundaries in raw speech signal</li> <li>Achieved highest boundary detection rate (6.5 % higher F-score) on TIMIT (Matlab, keras, kaldi)</li> </ul>
OTHER PROJECTS	<b>Parkinson's detection from Speech and accelerometer data</b> <span style="float: right;">Aug '18 - Mar '20</span> <i>Advisor: Dr. Najim Dehak, JHU</i> <ul style="list-style-type: none"> <li>Proposed a Siamese networks based method Parkinson detection from Speech data</li> <li>BeatPD challenge for detecting Dyskinesia, Termor from accelerometer data (Pytorch, kaldi)</li> </ul> <b>Instantaneous Frequency based Automatic Speech Recognition</b> <span style="float: right;">July '16 - Feb '17</span> <i>Advisor: Dr. K. Sri Rama Murty, IIT Hyderabad</i> <ul style="list-style-type: none"> <li>MFCC features are used for speech recognition which don't utilize phase information</li> <li>Combined the proposed IFCC features with traditional MFCC to improve ASR performance by 6% relative on TIMIT dataset (Matlab, kaldi, keras)</li> </ul>
SKILLS	<b>Programming Languages:</b> Python, bash <b>Tools/Frameworks:</b> Pytorch, Matlab, keras, kaldi, Tensorflow, Theano, SciKit, HTK
INVITED TALKS	Hidden Markov Models for Speech Recognition <ul style="list-style-type: none"> <li>TEQIP workshop, Rajiv Gandhi Institute of Technology, Kerala, 2017</li> </ul> Self-expressing autoencoders for unsupervised feature learning <ul style="list-style-type: none"> <li>University of Illinois at Urbana-Champaign, 2020</li> </ul> Segmental Contrastive Predictive Coding for Unsupervised Acoustic Segmentation <ul style="list-style-type: none"> <li>ISCA SIGML Seminar Series, 2021</li> </ul>
TEACHING EXPERIENCE	<b>Teaching Assistant</b> <ul style="list-style-type: none"> <li>EN.520.612: Machine learning for signal processing, JHU <span style="float: right;">Aug '19 - Dec '19</span>  <i>Introduces PCA, PPCA, ICA, NMF, GMMs, HMMs, DNNs, RNNs.</i></li> <li>EE7390: Pattern recognition and Machine learning, IIT Hyderabad <span style="float: right;">Aug '15 - Dec '15, Jan '17 - May '17</span>  <i>Introduces k-means, K-NN, GMMs, HMMs, DNNs, Linear and Fisher discriminators etc.</i></li> <li>EE5370: Introduction to Machine learning, IIT Hyderabad <span style="float: right;">Aug '16 - Dec '16</span>  <i>Introduces Naive Bayes classifier, Support vector machines, and clustering techniques.</i></li> </ul>
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