SOHEIL KHORRAM

soheil.khorram@utdallas.edu soheil-khorram.github.io soheil-khorram (734)-730-8731
CRSS Lab, Department of Engineering and Computer Science, University of Texas at Dallas, USA

RESEARCH INTERESTS

- Deep Learning (CNN, LSTM/GRU, Deformable CNN, GAN, Attention, Transformer)
- Speech Processing (Emotion Recognition/Affective Computing, Speech Separation/Enhancement, Speech Synthesis)
- Machine Learning, Pattern Recognition, Graphical Models (HMM, CRF)

EXPERIENCE

Research Associate at CRSS Lab, University of Texas at Dallas, Richardson, TX, USA *09/2018 – Present*
Supervisor: John H. L. Hansen

- Working on a GAN-based speech enhancement system for cochlear implant users (PyTorch).
- Developed causal and non-causal CNN-based speech enhancement systems with vanilla/spectral-subtraction/Wiener style structures in cochlear implant auditory space (Keras). `github.com/soheil-khorram/DNN-based-speech-enhancement`
- Proposed and implemented probabilistic permutation invariant training (PIT) that extends and improves the conventional PIT for DNN-based speech separation (TensorFlow and Kaldi). `github.com/soheil-khorram/Prob-PIT`

Postdoc at CHAI Lab, University of Michigan, Ann Arbor, MI, USA *12/2015 – 09/2018*
Supervisors: Emily Mower Provost, Melvin G. McInnis

- Built the best performing speech-based continuous emotion recognizer on AVEC 2016 and AVEC 2017 datasets using a new network: delayed sinc convolutional network (TensorFlow and Kaldi). `github.com/soheil-khorram/MDS-network`
- Studied large receptive field convolutional networks including dilated, downsampling/upsampling, hourglass and depth-wise separable networks for speech emotion recognition (Keras). `github.com/soheil-khorram/neural-network`
- Designed and organized the collection procedure of the PRIORI Emotion Dataset – a large-scale (42 hrs) conversational dataset of emotional speech passively-recorded from phone conversations of patients with bipolar disorder.
- Worked on progressive neural networks for transferring knowledge between different paralinguistic tasks (emotion, speaker and gender recognition) and between different datasets (IEMOCAP and MSP-IMPROV) (PyTorch, openSMILE).
- Proposed a new time warping algorithm, trainable time warping (TTW), that estimates the DTW averaging solution in linear time and space complexity using a new deformable convolutional kernel, shifted sinc kernel. `github.com/soheil-khorram/TTW`
- Developed an i-vector/SVM-based depression recognition system for patient with bipolar disorder (LIBSVM, Kaldi).

Visiting Ph.D. student at CSTR Lab, University of Edinburgh, Edinburgh, UK *07/2013 – 12/2013*
Supervisor: Simon King

- Introduced new acoustic modeling techniques leveraging maximum entropy and Gaussian conditional random field models for statistical parametric speech synthesis.
- Improved F0 modeling in HMM-based speech synthesis using a soft decision tree structure.

Ph.D. student at SPL Lab, Sharif University of Technology, Tehran, Iran *09/2009 – 01/2015*
Supervisor: Hossein Sameti

- Developed a high-quality HMM-based speech synthesis system with MOS of 4 using multi-space distribution hidden semi-Markov model, STRAIGHT Vocoder and global variance-based parameter generation for both English and Persian languages.
- Proposed a Gaussian Conditional Random Field-based acoustic modeling for statistical parametric speech synthesis.

Member of Technical Staff, ASR-Gooyesh Pardaz Company, Tehran, Iran *06/2008 – 10/2015*
Supervisor: Hadi Veisi

- Worked on both NLP and DSP parts of a unit-selection-based TTS system (Festival, Festvox, Speech Tools and SPTK).
- Built a real-time TTS system for mobile applications using clustergen synthesizer and CART classifiers.

EDUCATION

University of Michigan, Ann Arbor, MI, USA *12/2015 – 09/2018*
Postdoc in Computer Science and Engineering

University of Edinburgh, Edinburgh, UK *07/2013 – 12/2013*
Visiting Ph.D. student at the Centre for Speech Technology Research

Sharif University of Technology, Tehran, Iran *09/2009 – 01/2015*
Ph.D. in Computer Engineering, Artificial Intelligence (GPA: 19.14/20)

Sharif University of Technology, Tehran, Iran *09/2006 – 11/2008*
M.Sc. in Computer Engineering (GPA: 18.67/20)

Shahid Bahonar University, Kerman, Iran *09/2002 – 09/2006*
B.Sc. in Computer Engineering (GPA: 16.18/20)

LEADERSHIP AND AWARDS

- Managed 18 students in the task of labeling speaker/gender information of the UTD forensic speaker recognition project in 2019.
- Supervised 22 students of the University of Michigan in a 2-year period (07/2016–08/2018) for annotating emotion labels of the PRIORI Emotion Dataset. The collected dataset contains 42 hours of telephonic conversations of subjects with bipolar disorder.
- Received two conference travel awards from the University of Michigan in 2016.
- Received a scholarship from the ASR Gooyesh Pardaz Company to participate in a research program at the University of Edinburgh (CSTR lab) in 2013.
- Earned distinction for being qualified and competent project manager in ASR-Gooyesh Pardaz Company in 2012.
- Ranked first among Computer Architecture M.Sc. students of the Sharif University of Technology in 2009.
- Ranked second among Computer Hardware Engineering B.Sc. students of the Shahid Bahonar University in 2006.
- Ranked second among 8998 participants of the Iranian nationwide graduate school entrance exam in the field of Computer Engineering (Computer Architecture), in 2006.

TECHNICAL SKILLS

Programming:	Python, MATLAB, Java, C#, C++, and C
Software & Tools:	**Deep Learning Libraries:** TensorFlow, PyTorch, Keras, Flashlight, ArrayFire
	Speech Processing Tools: Kaldi, openSMILE, HTK, HTS, Festival, Festvox, Speech Tools, SPTK, STRAIGHT, WORLD, MSR Identity, Praat, LIBSVM

PUBLICATIONS

Journal Publications

1. **Khorram S.**, McInnis M., Provost E. M., “Jointly Aligning and Predicting Continuous Emotion Annotations”, IEEE Transactions on Affective Computing, 2019.
2. Taheri M., **Khorram S.**, Assi M., Sameti H., and Bijankhan M., “Designing and Recording a Speech Database for Persian TTS Systems”, Journal of Comparative Linguistic Researches, article 5, vol. 6, issue 12, page 69-84, 2017 [In Persian].
3. **Khorram S.**, Sameti H., King S., “Soft Context Clustering for F0 Modeling in HMM-Based Speech Synthesis”, EURASIP Journal on Advances in Signal Processing, 2015.
4. **Khorram S.**, Sameti H., King S., Drugman T., “Context-Dependent Acoustic Modeling Based on Hidden Maximum Entropy Model for Statistical Parametric Speech Synthesis”, EURASIP Journal on Audio, Speech, and Music Processing, 2014.
5. **Khorram S.**, Sameti H., Bahmaninezhad F., “Spectral Modeling Based on Gaussian Conditional Random Field for Statistical Parametric Speech Synthesis”, CSI Journal on Computer Science and Engineering, vol. 10, no. 2&4(b), pages 4759, 2012.

Conference Publications

1. Jafarlou S., **Khorram S.**, Kothapally V., Hansen J., ”Analyzing Large Receptive Field Convolutional Networks for Distant Speech Recognition”, ASRU, 2019.
2. Ghorbani S., **Khorram S.**, Hansen J., ”Domain Expansion in DNN-based Acoustic Models for Robust Speech Recognition”, ASRU, 2019.
3. Yousefi M., **Khorram S.**, Hansen J., “Probabilistic Permutation Invariant Training for Speech Separation”, Interspeech, 2019.
4. Nursadul M., **Khorram S.**, Hansen J., “Convolutional Neural Network-based Speech Enhancement for Cochlear Implant Recipients”, Interspeech, 2019.
5. **Khorram S.**, McInnis M., Provost E. M., “Trainable Time Warping: Aligning Time-Series in the Continuous-Time Domain”, ICASSP, 2019.
6. Zhang B., **Khorram S.**, Provost E. M., “Exploiting Acoustic and Lexical Properties of Phonemes to Recognize Valence from Speech”, ICASSP, 2019.
7. **Khorram S.**, Jaiswal M., Gideon J., McInnis M., Provost E. M., “The PRIORI Emotion Dataset: Linking Mood to Emotion Detected In-the-Wild”, Interspeech, 2018.
8. Aldeneh Z., **Khorram S.**, Dimitriadis D., Provost E. M., “Pooling Acoustic and Lexical Features for the Prediction of Valence”, 19th ACM International Conference on Multimodal Interaction (ICMI), 2017.
9. **Khorram S.**, Aldeneh Z., Dimitriadis D., McInnis M., Provost E. M., “Capturing Long-term Temporal Dependencies with Convolutional Networks for Continuous Emotion Recognition”, Interspeech 2017.
10. Gideon J., **Khorram S.**, Aldeneh Z., McInnis M., Provost E. M., “Progressive Neural Networks for Transfer Learning in Emotion Recognition”, Interspeech 2017.
11. Gideon J., Zhang B., Aldeneh Z., Kim Y., **Khorram S.**, Le D., Provost E. M., “Wild Wild Emotion: A Multimodal Ensemble Approach”, 18th ACM International Conference on Multimodal Interaction (ICMI), 2016.
12. **Khorram S.**, Gideon J., McInnis M., Provost E. M., “Recognition of Depression in Bipolar Disorder: Leveraging Cohort and Person-Specific Knowledge”, Interspeech, 2016.
13. **Khorram S.**, Sameti H., Bahmaninezhad F., “Context-Dependent Deterministic Plus Stochastic Model”, 12th IEEE International Conference on Signal Processing (ICSSP), 2014.

14. **Khorram S.**, Bahmaninezhad F., Sameti H., "Speech Synthesis Based on Gaussian Conditional Random Fields", International Symposium on Artificial Intelligence and Signal Processing (AISP), 2013.
15. Saleh F. S., Shams B., Sameti H., **Khorram S.**, "An Automatic Prosodic Event Detector Using MSD HMMs for Persian Language", International Symposium on Artificial Intelligence and Signal Processing (AISP), 2013.
16. Bahmaninezhad F., **Khorram S.**, Sameti H., "Average Voice Modeling Based on Unbiased Decision Trees", Advances in Nonlinear Speech Processing, 2013.
17. Bahmaninezhad F., Sameti H., **Khorram S.**, "HMM-Based Persian Speech Synthesis Using Limited Adaptation Data", 11th IEEE International Conference on Signal Processing (ICSP), 2012.
18. **Khorram S.**, Sameti H., Veisi H., "An Optimum MMSE Post-Filter for Adaptive Noise Cancellers in Automobile Environment", International Conference on Information Science, Signal Processing and Their Applications, 2012.
19. Bahaadini S., Sameti H., **Khorram S.**, "Implementation and Evaluation of Statistical Parametric Speech Synthesis Methods for the Persian Language", IEEE International Workshop on Machine Learning for Signal Processing (MLSP), 2011.
20. **Khorram S.**, Sameti H., Veisi H., "LP-Based Over-Sampled Subband Adaptive Noise Canceller for Speech Enhancement in Diffuse Noise Fields", IEEE 9th International Conference on Signal Processing (ICSP), 2008.
21. **Khorram S.**, Sameti H., Veisi H., Abutalebi H. R., "A New Lattice LP-based Post-filter for Adaptive Noise Cancellers in Mobile and Vehicular Applications", 8th IEEE Symposium on Signal Processing and Information Theory, 2008.

REFEREE REVIEWER

- IEEE Transactions on Affective Computing
- IEEE Transactions on Audio, Speech and Language Processing
- IEEE Transactions on Emerging Topics in Computational Intelligence
- ACM International Conference on Multimodal Interaction (ICMI)
- Springer - Language Resources and Evaluation Editors

REFERENCES

Will be available upon request.