Curriculum Vitae

Yu Tsao (曹昱)

Research Fellow (Professor)/ Deputy Director
Bio-Acoustic Signal Processing (Bio-ASP) Lab: http://bio-asplab.citi.sinica.edu.tw/
Research Center for Information Technology Innovation, Academia Sinica
128 Academia Rd., Sec. 2, Nankang District, Taipei 115, Taiwan
TEL: +886-22-2787-2390
Email: yu.tsao@citi.sinica.edu.tw http://www.citi.sinica.edu.tw/pages/yu.tsao/
Google Scholar: https://scholar.google.com/citations?hl=zh-TW&user=ZO5e5I4AAAAJ
Google Scholar Citation: 6663 (1247 until 2022/09/07).

RESEARCH INTERESTS

- Speech Enhancement and Voice Conversion
- Assistive Oral Communication Technologies
- Biomedical Acoustic Signal Processing
- Deep Learning Algorithm Development

EDUCATION

GEORGIA INSTITUTE OF TECHNOLOGY, Atlanta, Georgia
Ph.D. in Electrical and Computer Engineering Aug. 2003-Dec. 2008
- Research Topic: Robust Speech Recognition, under advisor: Dr. Chin-Hui Lee
- Leadership Activities: President, Taiwanese Student Association (TSA): represented the school; assisted students and visiting scholars from Taiwan.

NATIONAL TAIWAN UNIVERSITY, Taipei, Taiwan
Master of Science in Electrical Engineering Sept. 1999-June 2001
Bachelor of Science in Electrical Engineering Sept. 1995-June 1999
- Graduate Research Topic: Rapid Speaker Adaptation, under advisor Dr. Lin-Shan Lee
- Leadership Activities: Chairman, Public Relations Department of the Student Association: organized events to cultivate strong relationships among members.

PROFESSIONAL EXPERIENCE

RESEARCH CENTER FOR INFORMATION TECHNOLOGY INNOVATION, ACADEMIA SINICA, Taipei, Taiwan
Researcher Fellow (Professor)/Deputy Director 2020/08–Present
Associate Researcher Fellow (Associate Professor) 2016/05–2020/08
Assistant Researcher Fellow (Assistant Professor) 2011/11–2016/04
- Research and develop speech signal processing algorithms.
- Develop devices of assistive oral communication technology.
- Derive acoustic signal processing for biomedical applications.
- Derive novel machine learning algorithms for acoustic signal processing.

NATIONAL INSTITUTE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY, Kyoto, Japan
Expert Researcher April 2009-Sept. 2011
- Developed research to handle non-native accent issues for automatic speech recognition applications.
- Developed digital signal processing-based solutions to improve the performance of speech recognizers under real-world adverse conditions.
- Contributed to developing VoiceTra multilingual speech-to-speech translation application on iPhone.
- Carried out projects of field tests and dissemination of spoken dialog interface technologies.

DEVELOPED TOOLS

(1) A smartphone-based assistive listening system, **SmartHear**, uses wireless technologies for individuals with mild-to-moderate hearing loss. **SmartHear** has been developed as a **mobile application program** and is available on **Google Play** (since 2015, there have been more than **10,000 downloads**). A demo **video** about **SmartHear** can be viewed at https://www.youtube.com/watch?v=e9HqIj09QJs.
(2) We have open-sourced codes and datasets of our biomedical and speech signal processing experiments: https://bio-asplab.citi.sinica.edu.tw/Opensource.html

HONORS

- **Clarity Machine Learning Challenges for Hearing Aids Challenge Gold Prize (2022)**
- **IEEE Signal Processing Society (SPS) Young Author Best Paper Award (**corresponding author**) (2021)**
- National Innovation Award, Taiwan (2021)
- National Innovation Award, Taiwan (2020)
- National Innovation Award, Taiwan (2019)
- National Innovation Award, Taiwan (2018)
- Outstanding Elite Award, Chung Hwa Rotary Educational Foundation 2019-2020 (2019–present)
- Travel Grant, ICML (2019)
- Distinguished Lecture Award, APSIPA (2018)
- Best Student Paper Award, ISCSLP 2018 (2018)
- Poster Presentation Award, APSIPA 2017 (2017)
- Career Development Award, Academia Sinica, Taiwan (2017)

PROFESSIONAL ACTIVITIES

Tutorial Lecture

- Title: Generative Adversarial Network and its Applications to Speech Signal Processing and Natural Language Processing, ICASSP 2018 (2018/04)
- Title: Generative Adversarial Network and its Applications to Speech Signal Processing and Natural Language Processing, APSIPA 2018 (2018/11)
- Title: Generative Adversarial Network and its Applications to Speech Signal Processing and Natural Language Processing, ISCSLP 2018 (2018/11)
- Title: Generative Adversarial Network and its Applications to Speech Signal Processing and Natural Language Processing, Interspeech 2019 (2019/09)
- Title: Speech Enhancement based on Deep Learning and Intelligibility Evaluation, APSIPA 2019 (2019/11).
- Title: Speech Enhancement based on Deep Learning and Intelligibility Evaluation, Interspeech 2020 (2020/10)
- Title: Theory and Practice of Voice Conversion, APSIPA 2020 (2020/12).
- Title: Speech Perception and Enhancement in Cochlear Implants, APSIPA 2021 (2021/12).
- Title: Speech Perception and Enhancement in Cochlear Implants, EUSIPCO 2022 (2022/08).

Invited Talks

- Keynote Speech, IEEE ICTS 2021 (2021/10)
- Invited Lecture APSIPA Japan Chapter (2021/03)

Challenge Organizer

- Voice Detection Challenge in IEEE Big Data 2018. The challenge has attracted the participation of **109 teams from 27 different countries**.
- VoiceMOS Challenge, a potential special session in Interspeech 2022.

International Services

- Chair, Speech, Language, and Audio (SLA) Technical Committee, APSIPA (2020-2021)
- Distinguished Lecturer, APSIPA (2019-2021)

Editorship

- Associate Editor, IEEE Signal Processing Letters (2020-present)
- Associate Editor, IEEE/ACM Transactions on Audio, Speech and Language Processing (2019-present)
- Associate Editor, APSIPA Transactions on Signal and Information Processing (2022- present)
- Associate Editor, IEICE Transactions on Information and Systems (2016-2020)

Conference Services

- Winter School Lecturer, APSIPA 2019, Lanchou, China (2019/11)
- Session Chair, APSIPA 2019, Lanchou, China (2019/11)
- Special Session Chair, Interspeech 2019, Austria (2019/09)
- Session Chair, IALP 2019, Taipei, Taiwan (2019/08)
- Local Arrangement Chair, ISCSLP 2018, Taipei, Taiwan (2018/11)
- Session Chair, APSIPA 2018, Honolulu, USA (2018/11)
- Special Session Chair, APSIPA 2017, Kuala Lumpur, Malaysia (2017/12)
- Special Session Chair, ISCSLP 2016, Tianjin (2016/10)
- Session Chair, ROCLING 2015, Taiwan (2015/10)
- Special Session Chair, ICASSP 2015, South Brisbane, Queensland, Australia (2015/04)
- Session Chair, ROCLING 2014, National Central University, Taiwan (2014/09)
- Special Session Chair, APSIPA, Kaohsiung, Taiwan (2013/10)
- Session Chair, ROCLING 2013, National Sun Yat-sen University, Taiwan (2013/10)
- Session Chair, ROCLING 2012, Yuan Ze University, Taiwan (2012/09)
- Session Chair, ISCSLP 2010, National Cheng Kung University, Taiwan (2010/11)
- Session Chair, ROCLING 2012, Yuan Ze University, Taiwan (2012/09)
- Committee Members, 2012 IEEE SPS Summer School, National Tsing Hua University, Taiwan (2012/07)

Conference Chairs

- Conference Chair, ROCLING 2017, Taiwan (2017/11–2017/11)
- Sponsor Chair, TAAI 2017, Taiwan (2017/11–2017/12)
- Program Chair , ROCLING 2016, Taiwan (2016/10–2016/10)
- Organizer, SWS 2015, IIS, Academia Sinica, Taiwan (2015/03–2015/03)

REPRESENTATIVE RESEARCH WORKS

My research focus is to derive novel machine learning and speech signal processing algorithms for assisting people who have hearing and speaking disabilities, which can be divided into three parts: speech enhancement and voice conversion, assistive hearing technologies, and assistive speaking technologies.

Speech Enhancement and Voice Conversion

- Paper#1: Speech Enhancement Based on Deep Denoising Autoencoder, Interspeech 2013, **Google Citation: 818** (until 2022/09/07).
- Paper#2: Voice Conversion from Unaligned Corpora using Variational Autoencoding Wasserstein Generative Adversarial Networks, *Interspeech 2017*, **Google Citation: 369** (until 2022/09/07).
- Paper#3: Voice Conversion from Non-parallel Corpora using Variational Auto-encoder, *APSIPA 2016*, **Google Citation: 267** (until 2022/09/07).
- Paper#4: End-to-End Waveform Utterance Enhancement for Direct Evaluation Metrics Optimization by Fully Convolutional Neural Networks, *IEEE/ACM Transactions on Audio, Speech and Language Processing*, vol. 26(9), pp. 1570-1584, April 2018. (**2021 IEEE Signal Processing Society (SPS) Young Author Best Paper Award**). **Google Citation: 228** (until 2022/09/07).
- Paper#5: MetricGAN: Generative Adversarial Networks based Black-box Metric Scores Optimization for Speech Enhancement, in *Proc. ICML 2019*, **Long Oral Presentation** with **ICML Travel Grant**. **Google Citation: 140** (until 2022/09/07).
- Paper#6: Audio-Visual Speech Enhancement Using Multimodal Deep Convolutional Neural Networks, *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 2(2), pp. 117-128, April. 2018.

Google Citation: 173 (until 2022/09/07).

- Paper#7: Multichannel Speech Enhancement by Raw Waveform-mapping using Fully Convolutional Networks, *IEEE/ACM Transactions on Audio, Speech and Language Processing*, vol. 28, pp. 1888-1900, Feb. 2020. **Google Citation: 35** (until 2022/09/07).
- Paper#8: Unsupervised Noise Adaptive Speech Enhancement by Discriminator-Constrained Optimal Transport, *NeurIPS* 2021. **Google Citation: 2** (until 2022/09/07).

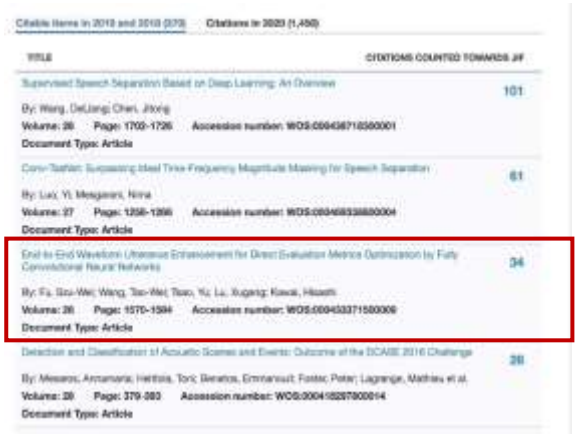
Assistive Hearing Technologies

- Paper#9: A Deep Denoising Autoencoder Approach to Improving the Intelligibility of Vocoded Speech in Cochlear Implant Simulation, *IEEE Transactions on Biomedical Engineering*, vol. 64(7), pp. 1568-1578, July, 2017. **Google Citation: 92** (until 2022/09/07).
- Paper#10: Deep learning–based noise reduction approach to improve speech intelligibility for cochlear implant recipients, Ear and Hearing, 2018.vol. 4, pp. 795-809, **Google Citation: 62** (until 2022/09/07).
- Paper#11: Improving the Intelligibility of Speech for Simulated Electric and Acoustic Stimulation Using Fully Convolutional Neural Networks*, IEEE Transactions on Neural Systems and Rehabilitation Engineering,* vol. 29, pp. 184-195, Dec. 2020. **Google Citation: 10** (until 2022/09/07).
- Paper#12: A Study of Joint Effect on Denoising Techniques and Visual Cues to Improve Speech Intelligibility in Cochlear Implant Simulation, *IEEE Transactions on Cognitive and Developmental*, vol. 13(4), pp. 984-994, Dec. 2021. **Google Citation: 6** (until 2022/09/07).

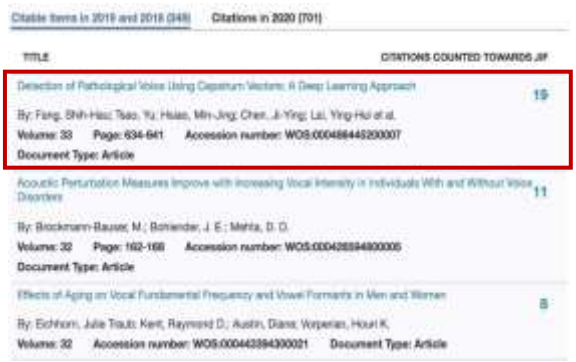
Assistive Speaking Technologies

- Paper#13: Joint Dictionary Learning-based Non-Negative Matrix Factorization for Voice Conversion to Improve Speech Intelligibility After Oral Surgery, *IEEE Transactions on Biomedical Engineering*, vol. 64 (11), pp. 2584-2594, November 2017. **Google Citation: 43** (until 2022/09/07).
- Paper#14: Generative Adversarial Networks for Unpaired Voice Transformation on Impaired Speech*, Interspeech 2018*. **Google Citation: 28** (until 2022/09/07).
- Paper#15: Detection of Pathological Voice Using Cepstrum Vectors: A Deep Learning Approach, *Journal of Voice*, vol 33(5), pp. 634-641, September 2019. **Google Citation: 152** (until 2022/09/07).

To date, my publications have received **6663 citations** based on **Google Scholar (1247 citations received in 2022)**. Among the publications mentioned above, Paper#4: End-to-End Waveform Utterance Enhancement for Direct Evaluation Metrics Optimization by Fully Convolutional Neural Networks) received the **ranking #3 citation** in IEEE/ACM Transactions on Audio, Speech and Language Processing 2020. Paper#15: Detection of Pathological Voice Using Cepstrum Vectors: A Deep Learning Approach, received the **ranking #1 citation** in Elsevier Journal of Voice 2020.



(a) IEEE/ACM Transactions on Audio, Speech, and Language Processing 2020 (reported on 2021/08/06)



(b) Elsevier Journal of Voice 2020 (reported on 2021/08/06)

FULL PUBLICATION LIST

(1) **Journal Papers (* indicating the corresponding author)**

[86] Y. Lin, **Y.Tsao**, and P.-J. Hsieh, "Neural Correlates of Individual Differences in Predicting Ambiguous Sounds Comprehension Level," to appear in NeuroImage.

[85] C.-H. Hu, Y.-H. Peng, J.Yamagishi, **Y. Tsao**, and H.-M. Wang, "SVSNet: An End-to-end Speaker Voice Similarity Assessment Model," to appear in IEEE Signal Processing Letters.

[84] S.-Y. Chuang, H.-M. Wang, and **Y. Tsao***, "Improved Lite Audio-Visual Speech Enhancement," to appear in IEEE Transactions on Audio, Speech and Language Processing.

[83] Y.-W. Chen, K.-H. Hung, Y.-J. Li, A. C.-F. Kang, Y.-S. Lai, K.-C. Liu, S.-W. Fu, S.-S. Wang, **Y. Tsao***, "CITISEN: A Deep Learning-Based Speech Signal-Processing Mobile Application," to appear in IEEE Access.

[82] K.-C. Liu, K.-H. Hung, C.-Y. Hsieh, H.-Y. Huang, C.-T. Chan, and Y. Tsao, "Deep Learning Based Signal Enhancement of Low-Resolution Accelerometer for Fall Detection Systems," to appear in IEEE Transactions on Cognitive and Developmental Systems.

[81] C.-T. Wang, Z.-Y. Chuang, C.-H. Hung, **Y. Tsao**, S.-H. Fang, "Detection of Glottic Neoplasm Based on Voice Signals Using Deep Neural Networks," to appear in IEEE Sensors Journal, (Letters)

[80] S.-S. Wang, C.-C. Lai, C.-T. Wang, **Y. Tsao**, S.-H. Fang, "Continuous Speech for Improved Learning Pathological Voice Disorders," to appear in IEEE Open Journal of Engineering in Medicine and Biology.

[79] Y.-C. Lin, C. Yu, Y.-T. Hsu, S.-W. Fu, **Y. Tsao***, T.-W. Kuo, "SEOFP-NET: Compression and Acceleration of Deep Neural Networks for Speech Enhancement Using Sign-Exponent-Only Floating-Points," to appear in IEEE Transactions on Audio, Speech and Language Processing.

[78] X. Lu, P. Shen, **Y. Tsao**, and H. Kawai, "Coupling A Generative Model With A Discriminative Learning Framework for Speaker Verification," to appear in IEEE Transactions on Audio, Speech and Language Processing.

[77] L. Chen, J.-T. Sheu, Y.-J. Chuang, K.-C. Liu, **Y. Tsao***, "Predicting the Travel Distance of Patients to Access Healthcare using Deep Neural Networks," to appear in IEEE Journal of Translational Engineering in Health and Medicine.

[76] K.-C. Liu, K.-H. Hung, C.-Y. Hsieh, H.-Y. Huang, C.-T. Chan, and **Y. Tsao***, "Deep Learning Based Signal Enhancement of Low-Resolution Accelerometer for Fall Detection Systems," to appear in IEEE Transactions on Cognitive and Developmental Systems.

[75] R.-Y. Tseng, T.-W. Wang, S.-W. Fu, C.-Y. Lee, and **Y. Tsao***, "A Study of Joint Effect on Denoising Techniques and Visual Cues to Improve Speech Intelligibility in Cochlear Implant Simulation," IEEE Transactions on Cognitive and Developmental Systems 2021.

[74] F. S. Abousaleh, W.-H. Cheng, N.-H. Yu, and **Y. Tsao***, "Multimodal Deep Learning Framework for Image Popularity Prediction on Social Media," IEEE Transactions on Cognitive and Developmental Systems, 2021.

[73] K.-C. Liu, M. Chan, C.-Y. Hsieh, H.-Y. Huang, C.-T. Chan, **Y. Tsao***, "Domain-adaptive Fall Detection Using Deep Adversarial Training," IEEE Transactions on Neural Systems & Rehabilitation Engineering, 2021.

[72] W. Ariyanti, T. Hussain, J.-C. Wang, C.-T. Wang, S.-H. Fang, and **Y. Tsao***, "Ensemble and Multimodal Learning for Pathological Voice Classification," IEEE Sensors Journal 2021.

[71] T.-H. Lin, T. Akamatsu, **Y. Tsao**, "Sensing ecosystem dynamics via audio source separation: A case study of marine soundscapes off northeastern Taiwan," PLOS Computational Biology 2021.

[70] H.-S. Lee, **Y. Tsao**, S.-K. Jeng, and H.-M. Wang, "Subspace-based Representation and Learning for Phonotactic Spoken Language Recognition," IEEE Transactions on Audio, Speech and Language Processing, 2021.

[69] N. Y.-H. Wang, H.-L. S. Wang, T.-W. Wang, S.-W. Fu, X. Lu, H.-M. Wang, and **Y. Tsao***," Improving the Intelligibility of Speech for Simulated Electric and Acoustic Stimulation Using Fully Convolutional Neural Networks," IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2020.

[68] T.-A. Hsieh, H.-M. Wang, X. Lu, and **Y. Tsao***, "WaveCRN: An Efficient Convolutional Recurrent Neural Network for End-to-end Speech Enhancement," IEEE Signal Processing Letters, 2020.

[67] K.-H. Tsai, W.-C. Wang, C.-H. Cheng, C.-Y. Tsai, J.-K. Wang, T.-H. Lin, S.-H. Fang, L.-C. Chen, and **Y. Tsao***, "Blind Monaural Source Separation on Heart and Lung Sounds Based on Periodic-Coded Deep Autoencoder," IEEE Journal of Biomedical and Health Informatics, 2020.

[66] W.-C. Huang, H. Luo, H.-T. Hwang, C.-C. Lo, Y.-H. Peng, **Y. Tsao***, and H.-M. Wang, "Unsupervised Representation Disentanglement using Cross Domain Features and Adversarial Learning in Variational Autoencoder based Voice Conversion," IEEE Transactions on Emerging Topics in Computational Intelligence, 2020.
[65] T. Hussain, S. M. Siniscalchi, H.-L. S. Wang, **Y. Tsao***, S. V. Mario, and W.-H. Liao, "Ensemble Hierarchical Extreme Learning Machine for Speech Dereverberation," IEEE Transactions on Cognitive and Developmental Systems, 2020.
[64] C. Yu, K.-H. Hung, S.-S. Wang, **Y. Tsao***, and J.-w. Hung, "Time-Domain Multi-modal Bone/air Conducted Speech Enhancement," IEEE Signal Processing Letters, 2020.
[63] S. C. Hidayati, T. W. Goh, Ji.-S. G. Chan, C.-C. Hsu, J. See, L.-K. Wong, K.-L. Hua, **Y. Tsao**, and W.-H. Cheng, "Dress With Style: Learning Style from Joint Deep Embedding of Clothing Styles and Body Shapes," IEEE Transactions on Multimedia, 2020.
[62] C. Yu, R. E. Zezario, S.-S. Wang, J. Sherman, Y.-Y. Hsieh, X. Lu, H.-M. Wang, and **Y. Tsao***, "Speech Enhancement based on Denoising Autoencoder with Multi-branched Encoders," IEEE Transactions on Audio, Speech and Language Processing, 2020.
[61] J.-K. Wang, Y.-F. Chang, K.-H. Tsai, W.-C. Wang, C.-Y. Tsai, C.-H. Cheng, and **Y. Tsao***, "Automatic Recognition of Murmurs of Ventricular Septal Defect using Convolutional Recurrent Neural Networks with Temporal Attentive Pooling," Scientific Reports, 2020
[60] C.-L. Liu, S.-W. Fu, Y.-J. Li, J.-W. Huang, H.-M. Wang, and **Y. Tsao***, "Multichannel Speech Enhancement by Raw Waveform-mapping using Fully Convolutional Networks," IEEE Transactions on Audio, Speech and Language Processing 2020.
[59] X. Wang et al.,, "ASVspoof 2019: ASVspoof 2019: A Large-scale Public Database of Synthetized, Converted and Replayed Speech," Computer Speech and Language, 2020.
[58] M. Lee, L. Lin, C.-Y. Chen, **Y. Tsao**, T.-H. Yao, M.-H. Fei and S.-H. Fang, "Forecasting Air Quality in Taiwan by Using Machine Learning," Scientific Reports, 2020.
[57] Y.-H. Lai, W.-N. Chen, T.-C. Hsu, C. Lin, **Y. Tsao**, and S. Wu, "Overall Survival Prediction of Non-small Cell Lung Cancer by Integrating Microarray and Clinical Data with Deep Learning," Scientific Reports, 2020.
[56] J.-Y. Wu, C. Yu, S.-W. Fu, C.-T. Liu, S.-Y. Chien, **Y. Tsao***, "Increasing Compactness of Deep Learning based Speech Enhancement Models with Parameter Pruning and Quantization Techniques," IEEE Signal Processing Letters, 2020.
[55] S.-W. Fu, C.-F. Liao, **Y. Tsao***, "Learning with Learned Loss Function: Speech Enhancement with Quality-Net to Improve Perceptual Evaluation of Speech Quality," IEEE Signal Processing Letters 2020.
[54] N. Y.-H. Wang, C.-H. Chiang, H.-L. S. Wang and **Y. Tsao***, "Atypical Frequency Sweep Processing in Chinese Children With Reading Difficulties: Evidence From Magnetoencephalography," 2020.
[53] C.-T. Wang, F.-C. Lin, J.-Y. Chen, M.-J. Hsiao, S.-H. Fang, Y.-H. Lai, **Y. Tsao**, "Detection of Pathological Voice Using Cepstrum Vectors: A Deep Learning Approach," Journal of Voice, 2019.
[52] S.-H. Fang, C.-T. Wang, J.-Y. Chen, **Y. Tsao** and F.-C. Lin, "Combining Acoustic Signals and Medical Records to Improve Pathological Voice Classification," APSIPA Transactions on Signal and Information Processing, 2019.
[51] T.-H. Lin and **Y. Tsao**, "Source Separation in Ecoacoustics: A Roadmap towards Versatile Soundscape Information Retrieval," Remote Sensing in Ecology and Conservation, 2019.
[50] C.-W. Lee et al., "Bioimaging: New Templated Ostwald Ripening Process of Mesostructured FeOOH for Third-Harmonic Generation Bioimaging," Small 2019.
[49] Y.-C. Chu, Y.-F. Cheng, Y.-H. Lai, **Y. Tsao**, T.-Y. Tu, S. T. Young, T.-S. Chen, Y.-F. Chung, F. Lai, W.-H. Liao, "A Mobile Phone–Based Approach for Hearing Screening of School-Age Children: Cross-Sectional Validation Study," JMIR Mhealth Uhealth, 2019.
[48] H.-T. Chiang, Y.-Y. Hsieh, S.-W. Fu, K.-H. Hung, **Y. Tsao***, S.-Y. Chien, "Noise Reduction in ECG Signals Using Fully Convolutional Denoising Autoencoders," IEEE Access, 2019
[47] **Y. Tsao***, T.-H. Lin, F. Chen, Y.-F. Chang, C.-H. Cheng, and K.-H. Tsai, "Robust S1 and S2 heart sound recognition based on spectral restoration and multi-style training," Biomedical Signal Processing and Control, 2019.
[46] H.-L. S. Wang , N. Y.-H. Wang , I-C. Chen, and **Y. Tsao***, "Auditory Identification of Frequency-Modulated Sweeps and Reading Difficulties in Chinese," Research in Developmental Disabilities, 2019.

[45] C.-T. Liu, T.-W. Lin, Y.-H. Wu, Y.-S. Lin, H. Lee, **Y. Tsao**, and S.-Y. Chien, "Computation-Performance Optimization of Convolutional Neural Networks with Redundant Filter Removal," IEEE Transactions on Circuits and Systems I, 2018.
[44] H.-P. Liu, **Y. Tsao***, and C.-S. Fuh, "Bone-Conducted Speech Enhancement Using Deep Denoising Autoencoder" Speech Communication 2018.
[43] **Y. Tsao***, H.-C. Chu, S.-H. Fang, J. Lee, and C.-M. Lin, "Adaptive Noise Cancellation using Deep Cerebellar Model Articulation Controller," IEEE Access, 2018.
[42] S.-W. Fu, T.-W. Wang, **Y. Tsao***, X. Lu, and H. Kawai, "End-to-End Waveform Utterance Enhancement for Direct Evaluation Metrics Optimization by Fully Convolutional Neural Networks," IEEE Transactions on Audio, Speech and Language Processing, 2018.
[41] T.-H. Lin, T. Akamatsu, and **Y. Tsao**, "Comparison of Passive Acoustic Soniferous Fish Monitoring with Supervised and Unsupervised Approaches," Journal of the Acoustical Society of America (JASA), 2018.
[40] J.-C. Hou, S.-S. Wang, Y.-H. Lai, **Y. Tsao***, H.-W. Chang, and H.-M. Wang, "Audio-visual Speech Enhancement based on Multimodal Deep Convolutional Neural Networks," IEEE Transactions on Emerging Topics in Computational Intelligence, 2018.
[39] S.-Y. Tsui, **Y. Tsao**, C.-W. Lin, S.-H. Fang, and C.-T. Wang, "Demographic and Symptomatic Features of Voice Disorders and Their Potential Application in Classification using Machine Learning Algorithms," Folia Phoniatrica et Logopaedica, 2018.
[38] S.-S. Wang, P. Lin, **Y. Tsao***, J.-W. Hung, and B. Su, "Suppression by Selecting Wavelets for Feature Compression in Distributed Speech Recognition," IEEE Transactions on Audio, Speech and Language Processing, 2018.
[37] Y.-H. Lai, **Y. Tsao**, X. Lu, F. Chen, Y.-T. Su, K.-C. Chen, Y.-H. Chen, L.-C. Chen, P.-H. Li, and C.-H. Lee, "Deep Learning based Noise Reduction Approach to Improve Speech Intelligibility for Cochlear Implant Recipients," Ear and Hearing, 2018.
[36] H.-T. Hwang, Y.-C. Wu, Y.-H. Peng, C.-C. Hsu, **Y. Tsao**, H.-M. Wang, Y.-R. Wang, and S.-H. Chen, "Voice Conversion based on Locally Linear Embedding," Journal of Information Science and Engineering, 2018.
[35] H.-T. Hwang, Y.-C. Wu, S.-S. Wang, C.-C. Hsu, **Y. Tsao**, H.-M. Wang, Y.-R. Wang, and S.-H. Chen, "Locally linear Embedding Based Post-filtering for Speech Enhancement," Journal of Information Science and Engineering, 2018.
[34] J. Torres-Sospedra et al., "Off-Line Evaluation of Mobile-Centric Indoor Positioning Systems: The Experiences from the 2017 IPIN Competition," Sensors, 2018.
[33] P. Lin, D. Lyu, F. Chen, S.-S. Wang, and **Y. Tsao***, "Multi-style Learning with Denoising Autoencoders for Acoustic Modeling in the Internet of Things (IoT)," Computer Speech and Language, 2017.
[32] T. Hussain, S. M. Siniscalchi, C.-C. Lee, S.-S. Wang, **Y. Tsao*** and W.-H. Liao, "Experimental Study on Extreme Learning Machine Applications for Speech Enhancement," IEEE Access, 2017.
[31] S.-H. Fang, Y.-X. Fei, Z. Xu, and **Y. Tsao**, "Learning Transportation Modes from Smartphone Sensors Based on Deep Neural Network," IEEE Sensors Journal, 2017.
[30] S.-W. Hsiao, H.-C. Sun, M.-C. Hsieh, M.-H. Tsai, **Y. Tsao**, and C.-C. Lee, "Toward Automating Oral Presentation Scoring during Principal Certification Program using Audio-Video Low-level Behavior Profiles," IEEE Transactions on Affective Computing, 2017.
[29] F. Chen, D. Zheng, and **Y. Tsao**, "Effects of Noise Suppression and Envelope Dynamic Range Compression on the Intelligibility of Vocoded Sentences for a Tonal Language," Journal of the Acoustical Society of America, 2017.
[28] X. Lu, P. Shen, **Y. Tsao**, and H. Kawai, "Regularization of Neural Network Model with Distance Metric Learning for I-vector based Spoken Language Identification," Computer Speech and Language, 2017.
[27] T.-H. Lin, S.-H. Fang, and **Y. Tsao***, "Improving Biodiversity Assessment via Unsupervised Separation of Biological Sounds from Long-duration Recordings," Scientific Reports, 2017.
[26] Y.-H. Lai, F. Chen, S.-S. Wang, X. Lu, **Y. Tsao***, and C.-H. Lee, "A Deep Denoising Autoencoder Approach to Improving the Intelligibility of Vocoded Speech in Cochlear Implant Simulation," IEEE Transactions on Biomedical Engineering, 2017.
[25] A. Chern, Y.-H. Lai, Y.-p. Chang, **Y. Tsao**, R. Y. Chang, and H.-W. Chang, "A Smartphone-Based Multi-Functional Hearing Assistive System to Facilitate Speech Recognition in the Classroom," IEEE Access, 2017,

This paper has been selected as a Featured Article (http://ieeeaccess.ieee.org/special-sections/featured-articles/smartphone-based-multi-functional-hearing-assistive-system-facilitate-speech-recognition-classroom/)
[24] T.-E. Chen, S.-I Yang, L.-T. Ho, K.-H. Tsai, Y.-H. Chen, Y.-F. Chang, Y.-H. Lai, S.-S. Wang, **Y. Tsao***, and C.-C. Wu, "S1 and S2 Heart Sound Recognition using Deep Neural Networks," IEEE Transactions on Biomedical Engineering, 2017.
[23] S.-W. Fu, P.-C. Li, Y.-H. Lai, C.-C. Yang, L.-C. Hsieh, and **Y. Tsao***, "Joint Dictionary Learning-based Non-Negative Matrix Factorization for Voice Conversion to Improve Speech Intelligibility After Oral Surgery," IEEE Transactions on Biomedical Engineering, 2016.
[22] H.-y. Lee, B.-H. Tseng, T.-H. Wen, and **Y. Tsao**, "Personalizing Recurrent Neural Network based Language Model by Social Network," IEEE Transactions on Audio, Speech and Language Processing, 2016.
[21] T. Guan, G.-x. Chu, **Y. Tsao**, F. Chen, "Assessing the Perceptual Contributions of Level-dependent Segments to Sentence Intelligibility," Journal of the Acoustical Society of America, 2016.
[20] S.-H. Fang, W.-H. Chang, **Y. Tsao**, H.-C. Shih, and C. Wang, "Channel State Reconstruction Using Multilevel Discrete Wavelet Transform for Improved Fingerprinting-Based Indoor Localization," IEEE Sensors Journal, 2016.
[19] H.-L. S. Wang, I-C. Chen, C.-H. Chiang, Y.-H. Lai, and **Y. Tsao**, "Auditory Perception, Suprasegmental Speech Processing, and Vocabulary Development in Chinese Preschoolers," Perceptual and Motor Skills, 2016.
[18] S.-H. Fang , H.-H. Liao , Y.-X. Fei , K.-H. Chen , J.- W. Huang , Y.-D. Lu and **Y. Tsao**, "Transportation Modes Classification Using Sensors on Smartphones," Sensors, 2016.
[17] S.-S. Wang, A. Chern, **Y. Tsao***, J.-w. Hung, X. Lu, Y.-H. Lai, B. Su, "Wavelet Speech Enhancement based on Nonnegative Matrix Factorization," IEEE Signal Processing Letters, 2016.
[16] F. Chen, **Y. Tsao**, and Y.-H. Lai, "Modeling Speech Intelligibility with Recovered Envelope from Temporal Fine Structure Stimulus," Speech Communication, 2016.
[15] P. Lin, S.-W. Fu, S.-S.Wang, Y.-H. Lai, and **Y. Tsao***, "Maximum Entropy Learning with Deep Belief Networks," Entropy, 2016.
[14] **Y. Tsao*** and Y.-H. Lai, "Generalized Maximum a Posteriori Spectral Amplitude Estimation for Speech Enhancement," Speech Communication, 2016.
[13] S.-H. Fang, C.-H. Wang, and **Y. Tsao**, "Compensating for Orientation Mismatch in Robust WiFi Localization Using Histogram Equalization," IEEE Transactions on Vehicular Technology, 2015.
[12] C.-C. Hsu, K.-M. Cheong, T.-S. Chi, and **Y. Tsao**, "Robust Voice Activity Detection Algorithm Based on Feature of Frequency Modulation of Harmonics and Its DSP Implementation," IEICE Transactions on Information and Systems, 2015.
[11] Y.-C. Lin, Y.-H. Lai, H.-W. Chang, **Y. Tsao**, Y.-p. Chang, and R. Y. Chang,, "A Smartphone-Based Remote Microphone Hearing Assistive System Using Wireless Technologies," IEEE Systems Journal, 2015, Smarthear Demo: https://www.youtube.com/watch?v=e9HqIj09QJs
[10] Y.-J, Lee, Y.-R. Chien, and **Y. Tsao**, "Rapid Converging M-max Partial Update Least Mean Square Algorithms with New Variable Step-size Methods," IEICE Transaction on Communications, 2015.
[9] Y.-H. Lai, **Y. Tsao**, F. Chen, "Effects of Adaptation Rate and Noise Suppression on the Intelligibility of Compressed-Envelope Based Speech," PLoS ONE, vol. 10.1371, journal.pone.0133519, July 2015.
[8] **Y. Tsao***, P. Lin, T.-y. Hu, and X. Lu, "Ensemble Environment Modeling using Affine Transform Group," Speech Communication, 2015.
[7] **Y. Tsao***, S.-H. Fang, and Y. Hsiao, "Acoustic Echo Cancellation Using a Vector-Space-Based Adaptive Filtering Algorithm," IEEE Signal Processing Letters, 2015.
[6] **Y. Tsao***, T.-y. Hu, S. Sakti, S. Nakamura, and L.-s. Lee, "Variable Selection Linear Regression for Robust Speech Recognition," IEICE Transactions on Information and Systems, 2014.
[5] **Y. Tsao***, X. Lu, P. Dixon, T.-y. Hu, S. Matsuda, and C. Hori, "Incorporating Local Information of the Acoustic Environments to MAP-based Feature Compensation and Acoustic Model Adaptation," Computer Speech and Language, 2014.
[4] **Y. Tsao**, S. Matsuda, C. Hori, H. Kashioka, and C.-H. Lee, "A MAP-based Online Estimation Approach to Ensemble Speaker and Speaking Environment," IEEE Transactions on Audio, Speech and Language Processing, 2014.
[3] Y.-H. Lai, **Y. Tsao***, and F. Chen, "A Study of Adaptive WDRC in Hearing Aids under Noisy Conditions,," International Journal of Speech & Language Pathology and Audiology, 2013, (invited paper).

[2] **Y. Tsao** and C.-H. Lee, "An Ensemble Speaker and Speaking Environment Modeling Approach to Robust Speech Recognition," IEEE Transactions on Audio, Speech and Language Processing, 2009.
[1] **Y. Tsao**, S.-M. Lee, and L.-S. Lee, "Segmental Eigenvoice with Delicate Eigenspace for Improved Speaker Adaptation," IEEE Transactions on Speech and Audio Processing, 2005.

(2) Book Chapter
[2] Y.-H. Lai, Fe. Chen, and **Y. Tsao**, "Adaptive Dynamic Range Compression for Improving Envelope-based Speech Perception," Emerging Technology and Architecture for Big-data Analytics, Anupam Chattopadhyay and Yu Hao, editor, pages 1, Springer,.
[1] P. Lin, **Y. Tsao**, and L.-W. Kuo, "Controlling the Biocompatibility and Mechanical Effects of Implantable Microelectrodes to Improve Chronic Neural Recordings in the Auditory Nervous System," An Excursus into Hearing Loss, IntechOpen

(3) <u>Selected International Conference Papers</u>

<u>Machine learning</u>
[5] C.-J. Hsu, H.-y. Lee, **Y. Tsao**, "XDBERT: Distilling Visual Information to BERT via Cross-Modal Encoders to Improve Language Understanding," to appear in ACL 2022.
[4] H.-Y. Lin, H.-H. Tseng, X. Lu, **Yu Tsao**, "Unsupervised Noise Adaptive Speech Enhancement by Discriminator-Constrained Optimal Transport," in Proc. NeurIPS 2021
[3] S.-W. Fu, C.-F. Liao, **Y. Tsao**, S.-D. Lin, "MetricGAN: Generative Adversarial Networks based Black-box Metric Scores Optimization for Speech Enhancement," in Proc. ICML 2018, **Long Oral** with **ICML Travel Grant Award** (**Acceptance rate=3%**).
[2] H. Jing, A.-C. Liang, S.-D. Lin, and **Y. Tsao**, "A Transfer Probabilistic Collective Factorization Model to Handle Sparse Data in Collaborative Filtering," in Proc. ICDM 2014 (**Acceptance rate=9.5%**)
[1] H. Jing, **Y. Tsao**, K.-Y. Chen and H.-M. Wang, "Semantic Naive Bayes Classifier for Document Classification," in Proc. IJCNLP 2013.

<u>Acoustic modeling and speech processing</u>
[73] X. Chang et al., "An Exploration of Self-Supervised Pretrained Representations for End-to-End Speech Recognition," in Proc. ASRU 2021.
[72] M. E Noor et al., "Investigation of A Single-Channel Frequency-Domain Speech Enhancement Network to Improve End-To-End Bengali Automatic Speech Recognition Under Unseen Noisy Conditions," in Proc. Oriental COCOSDA 2021.
[71] Y.-J. Lu, **Y. Tsao**, and S. Watanabe, "A Study on Speech Enhancement Based on Diffusion Probabilistic Model," in Proc. APSIPA ASC 2021
[70] X. Lu, P. Shen, **Y. Tsao**, and H. Kawai, "Siamese Neural Network with Joint Bayesian Model Structure for Speaker Verification," in Proc. APSIPA ASC 2021.
[69] Y.-J. Li, S.-S. Wang, **Y. Tsao**, and B. Su, "MIMO Speech Compression and Enhancement Based on Convolutional Denoising Autoencoder," in Proc. APSIPA ASC 2021.
[68] Y,-C. Wu, C.-H. Hu, H.-S. Lee, Y.-H. Peng, W.-C. Huang, **Y. Tsao**, H.-M. Wang and T. Toda, "Relational Data Selection for Data Augmentation of Speaker-dependent Multi-band MelGAN Vocoder," in Proc. Interspeech 2021
[67] S.-W. Fu, C. Yu, T.-A. Hsieh, P. Plantinga, M. Ravanelli, X. Lu, **Y. Tsao**, "MetricGAN +: An Improved Version of MetricGAN for Speech Enhancement," in Proc. Interspeech 2021.
[66] T.-A. Hsieh, C. Yu, S.-W. Fu, X. Lu, and **Y. Tsao**, "Improving Perceptual Quality by Phone-Fortified Perceptual Loss using Wasserstein Distance for Speech Enhancement," in Proc. Interspeech 2021.
[65] R. E Zezario, C.-S. Fuh, H.-M. Wang, **Y. Tsao**, "Speech Enhancement with Zero-Shot Model Selection," in Proc. EUSIPCO 2021
[64] G.-X. Lin, S.-W. Hu, Y.-J. Lu, **Y. Tsao**, and C.-S. Lu, "QISTA-Net-Audio: Audio Super-resolution via Non-Convex Lq-norm Minimization," in Proc. Interspeech 2021
[63] Y.-K. Wu, K.-P. Huang, **Y. Tsao**, H.-y. Lee, "One shot learning for speech separation," in Proc. ICASSP 2021.

[62] X. Lu, P. Shen, **Y. Tsao**, H. Kawai, "Unsupervised neural adaptation model based on optimal transport for spoken language identification," in Proc. ICASSP 2021
[61] C.-J. Peng, Y.-J. Chan, C. Yu, S.-S. Wang, **Y. Tsao**, T.-S. Chi, "Attention-based multi-task learning for speech-enhancement and speaker-identification in multi-speaker dialogue scenario," in Proc. ISCAS 2021.
[60] Y.-T. Chang, Y.-H. Yang, Y.-H. Peng, S.-S. Wang, T.-S. Chi, **Y. Tsao** and H.-M. Wang, "MoEVC: A Mixture of Experts Voice Conversion System With Sparse Gating Mechanism for Online Computation Acceleration," in Proc. ISCSLP 2021.
[59] R. E. Zezario, S.-W. Fu, C.-S. Fuh, **Y. Tsao**, and H.-M. Wang, "STOI-Net: A Deep Learning based Non-Intrusive Speech Intelligibility Assessment Model," in Proc. APSIPA ASC 2020
[58] S.-W. Fu et al., "Boosting Objective Scores of Speech Enhancement Model through MetricGAN Post-Processing," in Proc. APSIPA ASC 2020.
[57] H. Li, S.-W. Fu, **Y. Tsao**, J. Yamagishi, "iMetricGAN: Intelligibility Enhancement for Speech-in-Noise using Generative Adversarial Network-based Metric Learning," in Proc. Interspeech 2020.
[56] Y.-J. Lu, C.-F. Liao, X. Lu, J.-w. Hung, **Y. Tsao**, "Incorporating Broad Phonetic Information for Speech Enhancement," in Proc. Interspeech 2020 (**with ISCA Travel Grant Award**).
[55] S.-Y. Chuang, **Y. Tsao**, C.-C. Lo, H.-M. Wang, "Lite Audio-Visual Speech Enhancement," in Proc. Interspeech 2020 (**with ISCA Travel Grant Award**).
[54] C.-C. Lee, Y.-C. Lin, H.-T. Lin, H.-M. Wang, **Y. Tsao**, "SERIL: Noise Adaptive Speech Enhancement using Regularization-based Incremental Learning," in Proc. Interspeech 2020.
[53] R. E. Zezario, T. Hussain, X. Lu, H.-M. Wang, and **Y. Tsao**, "Self-supervised Denoising Autoencoder with Linear Regression Decoder for Speech Enhancement," in Proc. ICASSP 2020.
[52] P.-T. Huang, H.-S. Lee, S.-S. Wang, K.-Y. Chen, **Y. Tsao**, and H.-M. Wang, "Exploring the Encoder Layers of Discriminative Autoencoders for LVCSR," in Proc. Interspeech 2019, (**with ISCA Travel Grant Award**).
[51] C.-F. Liao, Y. Tsao, X. Lu and H. Kawai, "Incorporating Symbolic Sequential Modeling for Speech Enhancement," in Proc. Interspeech 2019, (**with ISCA Travel Grant Award**).
[50] W.-C. Huang, Y.-C. Wu, C.-C. Lo, P. L. Tobing, T. Hayashi, K. Kobayashi, T. Toda, **Y. Tsao** and H.-M. Wang, "Investigation of F0 conditioning and Fully Convolutional Networks in Variational Autoencoder based Voice Conversion," in Proc. Interspeech 2019, (**with ISCA Travel Grant Award)**.
[49] C.-F. Liao, Y. Tsao, H.-y. Lee and H.-M. Wang, "Noise Adaptive Speech Enhancement using Domain Adversarial Training," in Proc. Interspeech 2019 (**with ISCA Travel Grant**).
[48] W.-C. Huang, Y.-C. Wu, H.-T. Hwang, P. L. Tobing, T. Hayashiy, K. Kobayashi, T. Toda, **Y. Tsao**, H.-M. Wang, "Refined WaveNet Vocoder for Variational Autoencoder Based Voice Conversion," in Proc. EUSIPCO 2019
[47] Y.-C. Lin, Y.-T. Hsu, S.-W. Fu, **Y. Tsao**, and T.-W. Kuo, "IA-NET: Acceleration and Compression of Speech Enhancement using Integer-adder Deep Neural Network," in Proc. Interspeech 2019.
[46] L.-W. Chen, H.-Y. Lee, and **Y. Tsao**, "Generative Adversarial Networks for Unpaired Voice Transformation on Impaired Speech," in Proc. Interspeech 2019.
[45] X. Lu, P. Shen, S. Li, **Y. Tsao**, and H. Kawai, "Class-wise Centroid Distance Metric Learning for Acoustic Event Detection," in Proc. Interspeech 2019.
[44] F.-K. Chuang, S.-S. Wang, J.-w. Hung, **Y. Tsao**, and S.-H. Fang, "Speaker-aware Deep Denoising Autoencoder with Embedded Speaker Identity for Speech Enhancement," in Proc. Interspeech 2019.
[43] R. E. Zezario, S.-W. Fu, X. Lu, H.-M. Wang, and **Y. Tsao**, "Specialized Speech Enhancement Model Selection Based on Learned Non-Intrusive Quality Assessment Metric," in Proc. Interspeech 2019.
[42] T. Hussain, **Y. Tsao**, H.-M. Wang, J.-C. Wang, S. M. Siniscalchi, W.-H. Liao, "Audio-Visual Speech Enhancement Using Hierarchical Extreme Learning Machine," in Proc. EUSIPCO 2019.
[41] T. Hussain, **Y. Tsao**, S. M. Sinicalchi, J.-C. Wang, H.-M. Wang, and W.-H. Liao, "Bone-conducted Speech Enhancement using Hierarchical Extreme Learning Machine," in Proc. IWSDS 2019.
[40] C.-C. Lo, S.-w. Fu, W. C. Huang, X. Wang, J. Yamagishi, **Y. Tsao** and H.-M. Wang, "MOSNet: Deep Learning based Objective Assessment for Voice Conversion," in Proc. Interspeech 2019.
[39] K.-Y. Liu, S.-k. Lee, S.-S. Wang, **Y. Tsao**, J.-w. Hung, "Reducing noise and reverberation in speech signals via the integration of denoising autoencoder and temporal lowpass filtering," in Proc. ICASI 2019.

[38] W.-C. Huang et al., "Generalization of Spectrum Differential based Direct Waveform Modification for Voice Conversion," in Proc. ISCA SSW 10,
[37] Y.-L. Shen, C.-Y. Huang, S.-S. Wang, **Y. Tsao**, H.-M. Wang, and T.-S. Chi, "Reinforcement Learning Based Speech Enhancement for Robust Speech Recognition," in Proc. ICASSP 2019.
[36] S.-k. Lee, S.-S. Wang, **Y. Tsao**, J.-w. Hung, "Speech Enhancement Based on Reducing the Detail Portion of Speech Spectrograms in Modulation Domain via Discrete Wavelet Transform," in Proc. ISCSLP 2018.
[35] W.-C. Huang, H.-T. Hwang, Y.-H. Peng, **Y. Tsao**, H.-M. Wang, "Voice Conversion Based on Cross-Domain Features Using Variational Auto Encoders," in Proc. ISCSLP 2018 (**Best Student Paper Award**).
[34] X. Lu, P. Shen, S. Li, **Y. Tsao**, H. Kawai, "Temporal Attentive Pooling for Acoustic Event Detection," in Proc. Interspeech 2018.
[33] S.-W. Fu, **Y. Tsao**, H.-T. Hwang, H.-M. Wang, "Quality-Net: An End-to-End Non-intrusive Speech Quality Assessment Model based on BLSTM," in Proc. Interspeech 2018.
[32] Y.-H. Peng, H.-T. Hwang, Y.-C. Wu, **Y. Tsao**, H.-M. Wang, "Exemplar-Based Spectral Detail Compensation for Voice Conversion," in Proc. Interspeech 2018.
[31] L. Sun, J. Du, T. Gao, Y.-D. Lu, **Y. Tsao**, C.-H. Lee, N. Ryant, "A Novel LSTM-based Speech Preprocessor For Speaker Diarization in Realistic Mismatch Conditions," in Proc. ICASSP 2018.
[30] W.-J. Lee, S.-S. Wang, F. Chen, X. Lu, S.-Y. Chien, and **Y. Tsao**, "Speech Dereverberation Based on Integrated Deep and Ensemble Learning Algorithm," in Proc. ICASSP 2018.
[29] N. Ryant et al., "Enhancement and Analysis of Conversational Speech: JSALT 2017," ICASSP 2018.
[28] Y.-H. Peng, C.-C. Hsu, Y.-C. Wu, H.-T. Hwang, Y.-W. Liu, **Y. Tsao**, and H.-M. Wang, "Fast Locally Linear Embedding Algorithm for Exemplar-based Voice Conversion," in Proc. APSIPA ASC 2017 (**Poster Presentation Award**).
[27] S.-W. Fu, **Y. Tsao**, X. Lu, and H. Kawai, "Raw Waveform-based Speech Enhancement by Fully Convolutional Networks," in Proc. APSIPA ASC 2017.
[26] S.-W. Fu, T.-y. Hu, **Y. Tsao**, X. Lu, "Complex Spectrogram Enhancement by Convolutional Neural Network with Multi-metrics Learning," in Proc. MLSP 2017.
[25] C.-C. Hsu, H.-T. Hwang, Y.-C. Wu, **Y. Tsao**, and H.-M. Wang, "Voice Conversion from Unaligned Corpora Using Variational Autoencoding Wasserstein Generative Adversarial Networks," Interspeech2017.
[24] M.-H. Yang, H.-S. Lee, Y.-D. Lu, K.-Y. Chen, **Y. Tsao**, B. Chen, and H.-M. Wang, "Discriminative Autoencoders for Acoustic Modeling," in Proc. 2017.
[23] Y.-C. Wu, H.-T. Hwang, S.-S. Wang, C.-C. Hsu, **Y. Tsao**, and H.-M. Wang, "A Post-filtering Approach Based on Locally Linear Embedding Difference Compensation for Speech Enhancement," in Proc. Interspeech2017.
[22] C.-L. Wu, H.-P. Hsu, S.-S. Wang, J.-W. Hung, Y.-H. Lai, H.-M. Wang, and **Y. Tsao**, "Wavelet Speech Enhancement Based on Robust Principal Component Analysis," in Proc. Interspeech2017.
[21] J.-C. Hou, S.-S. Wang, Y.-H. Lai, J.-C. Lin, **Y. Tsao**, H.-W. Chang, and H.-M. Wang, "Audio-visual Speech Enhancement using Deep Neural Networks," in Proc. APSIPA ASC 2016.
[20] S.-W. Fu, **Y. Tsao**, X. Lu, "SNR-aware Convolutional Neural Network Modeling for Speech Enhancement," in Proc. Interspeech 2016..
[19] S.-S. Wang, H.-T. Hwang, Y.-H. Lai, **Y. Tsao**, X. Lu, H.-M. Wang, and B. Su, "Improving Denoising Auto-encoder based Speech Enhancement with the Speech Parameter Generation Algorithm," in Proc. APSIPA ASC 2015.
[18] H.-T. Hwang, **Y. Tsao**, H.-M. Wang, Y.-R. Wang, and S.-H. Chen, "A Probabilistic Interpretation for Artificial Neural Network-based Voice Conversion," in Proc. APSIPA ASC 2015.
[17] P. Lin, D.-C. Lyu, Y.-F. Chang, and **Y. Tsao**, "Temporal Alignment for Deep Neural Networks" in Proc. GlobalSIP 2015.
[16] X. Lu, P. Shen, **Y. Tsao**, C. Hori, H. Kawai, "Sparse Representation with Temporal Max-smoothing for Acoustic Event Detection," in Proc. Interspeech 2015.
[15] P. Lin, D.-C. Lyu, Y.-F. Chang, and **Y. Tsao**, "Speech Recognition with Temporal Neural Networks," in Proc. Interspeech2015.
[14] W.-C. Chen, P.-T. Lai, **Y. Tsao**, and C.-C. Lee, "Multimodal Arousal Rating using Unsupervised Fusion Technique," in Proc. ICASSP 2015.

[13] Y.-H. Lai, S.-S. Wang, P.-C. Li, and **Y. Tsao**, "A Discriminative Post-filter for Speech Enhancement in Hearing Aids," in Proc. ICASSP 2015.
[12] H.-S. Lee, **Y. Tsao**, H.-M. Wang and S.-K. Jen, "Clustering-based I-vector Formulation for Speaker Recognition," in Proc. Interspeech 2014.
[11] H. Jing, T.-Y. Hu, H.-S. Lee, W.-C. Chen, C.-C. Lee, **Y. Tsao** and H.-M. Wang, "Ensemble of Machine Learning Algorithms for Cognitive and Physical Speaker Load Detection," in Proc. Interspeech 2014.
[10] P. Lin, F. Chen, S.-S. Wang, **Y. Tsao**, and Y. H. Lai, "Automatic Speech Recognition with Primarily Temporal Envelope Information," in Proc. Interspeech 2014.
[9] X, Lu, **Y, Tsao**, S, Matsuda and C, Hori, "Ensemble Modeling of Denoising Autoencoder for Speech Spectrum Restoration," in Proc. Interspeech 2014.
[8] X. Lu, **Y. Tsao**, S. Matsuda, C. Hori, "Sparse Representation based on a Bag of Spectral Exemplars for Acoustic Event Detection," in Proc. ICASSP 2014.
[7] H.-t. Fan, J.-w. Hung, X. Lu, S.-S. Wang, **Y. Tsao**, "Speech Enhancement using Segmental Nonnegative Matrix Factorization," in Proc. ICASSP 2014.
[6] H.-T. Hwang, **Y. Tsao**, H.-M. Wang, Y.-R. Wang and S.-H. Chen, "Alleviating the Over-smoothing Problem in GMM-based Voice Conversion with Discriminative Training," in Proc. Interspeech 2013.
[5] T.-H. Wen, A. Heidel, H.-y. Lee, **Y. Tsao** and L.-S. Lee, "Recurrent Neural Network based Language Model Personalization by Social Network Crowdsourcing," in Proc. Interspeech 2013.
[4] X. Lu, **Y. Tsao**, S. Matsuda and C. Hori, "Speech Enhancement based on Deep Denoising Autoencoder," in Proc. Interspeech 2013.
[3] B. Li, **Y. Tsao** and K. C. Sim, "An Investigation of Spectral Restoration Algorithms for Deep Neural Networks based Noise Robust Speech Recognition," in Proc. Interspeech 2013.
[2] S.-S. Wang, **Y. Tsao**, J.-W. Hung, "Filtering on the Temporal Probability Sequence in Histogram Equalization for Robust Speech Recognition," in Proc. ICASSP 2013.
[1] Y.-C. Su, **Y. Tsao**, J.-E. Wu, F.-R. Jean, "Speech Enhancement using Generalized Maximum a Posteriori Spectral Amplitude Estimator," in Proc. ICASSP 2013.

Assistive listening technologies

[10] Z. Feng, **Yu Tsao**, and F. Chen, "Estimation and Correction of Relative Transfer Function for Binaural Speech Separation Networks to Preserve Spatial Cues," in Proc. APSIPA ASC 2021.
[9] Y.-W. Chen, K.-H. Hung, S.-Y. Chuang, J. Sherman, X. Lu, **Y. Tsao**, "A Study of Incorporating Articulatory Movement Information in Speech Enhancement," in Proc. EUSIPCO 2021.
[8] T.-Y. Lu, K.-C. Liu, Y.-W. Chen, K.-H. Hung, S.-Y. Chuang, J. Sherman, W.-C. Huang, X. Lu, **Y. Tsao**, "EMA2S: An End-to-End Multimodal Articulatory-to-Speech System," in Proc. ISCAS 2021.
[7] H.-T. Chiang, Y.-C. Wu, C. Yu, T. Toda, H.-M. Wang, Y.-C. Hu, and **Y. Tsao**, "HASA-NET: A Non-Intrusive Hearing-Aid Speech Assessment Network," in Proc. ASRU 2021.
[6] Y.-H. Lai, W.-Z. Zheng, S.-T. Tang, S.-H. Fang, W.-H. Liao, and **Y. Tsao**, "Improving the Performance of Hearing Aids in Noisy Environments based on Deep Learning Technology," in Proc. EMBC 2018.
[5] S.-S. Wang, **Y. Tsao**, H.-L. S. Wang, Y.-H. Lai, and L. P.-H. Li, "A Deep Learning based Noise Reduction Approach to Improve Speech Intelligibility for Cochlear Implant Recipients in the Presence of Competing Speech Noise," in Proc. APSIPA ASC 2017.
[4] Y.-T. Liu, R. Y. Chang, **Y. Tsao**, and Y.-p. Chang, "A New Frequency Lowering Technique for Mandarin-speaking Hearing Aid Users," in Proc. GlobalSIP 2015.
[3] Y.-T. Liu, **Y. Tsao**, and R. Y. Chang, "A Deep Neural Network based Approach to Mandarin Consonant/vowel Separation," in Proc. ICCE 2015.
[2] Y.-H. Lai, F. Chen, and **Y. Tsao**, "Effect of Adaptive Envelope Compression in Simulated Electric Hearing in Reverberation," in Proc. ISIC 2014.
[1] Y. H. Lai, F. Chen, and **Y. Tsao**, "An Adaptive Envelope Compression Strategy for Speech Processing in Cochlear Implants," in Proc. Interspeech 2014,

Assistive speaking technologies

[4] W.-C. Huang et al., "A Preliminary Study of a Two-Stage Paradigm for Preserving Speaker Identity in Dysarthric Voice Conversion," in Proc. Interspeech 2021.

[3] Y.-S. Liou et al., "Time Alignment Using Lip Images for Frame-Based Electrolaryngeal Voice Conversion," in Proc. APSIPA ASC 2021.
[2] M.-C. Yen et al., "Mandarin Electrolaryngeal Speech Voice Conversion with Sequence-to-Sequence Modeling," in Proc. ASRU 2021.
[1] C.-Y. Chen et al., "Enhancing Intelligibility of Dysarthric Speech Using Gated Convolutional-based Voice Conversion System," in Proc. Interspeech 2020.

(4) Patents

[3] **Y. Tsao**, S.-H. Fang, Y.-R. Chien, and Y. Shiao, "Acoustic echo cancellation method and system using prior knowledge of the acoustic conditions," US patent 2017.

[2] K.-H. Tsai, **Y. Tsao**, S.-H. Ku, T. Liang, Y.-F. Chang, S.-I. Yang, "Method and system for recognizing physiological sound," US Patent 2017.

[1] **Y. Tsao**, and S.-S. Wang, "System and methods for suppression by selecting wavelets for feature compression in distributed speech recognition," US Patent 2021.

[2] K.-H. Tsai, **Y. Tsao**, S.-H. Ku, T. Liang, Y.-F. Chang, S.-I. Yang, "Method and system for recognizing physiological sound" EP 3100675 A1 2018