

# Publications by Zhao Ren

Update time: Jan. 2026

## I. BOOK CHAPTERS

- 1) N. Cummins, Z. Ren, A. Mallol-Ragolta, and B. W. Schuller, *Artificial Intelligence in Precision Health*, ch. Chapter 5 – Machine learning in digital health, recent trends, and ongoing challenges, pp. 121–148. Elsevier, 2020

## II. GUEST EDITORIALS

- 2) Z. Ren, B. W. Schuller, B. M. Eskofier, T. N. Nguyen, and W. Nejdl, “Guest editorial trustworthy and collaborative AI for personalised healthcare through edge-of-things,” *IEEE Journal of Biomedical and Health Informatics*, pp. 5213 – 5215, Nov. 2023

## III. JOURNAL ARTICLES

- 3) T. T. Nguyen, Z. Ren, T. Pham, P. Le Nguyen, Q. V. H. Nguyen, and H. Yin, “A review of instruction-guided image editing,” *Engineering Applications of Artificial Intelligence*, vol. 163, p. 112953, 2026
- 4) T. T. Nguyen, T. T. Huynh, Z. Ren, P. L. Nguyen, A. W.-C. Liew, H. Yin, and Q. V. H. Nguyen, “A survey of machine unlearning,” *ACM Transactions on Intelligent Systems and Technology*, vol. 16, no. 5, 2025. 46 pages
- 5) P. Wißbrock, L. Koschek, Z. Ren, and W. Nejdl, “Enhancing quality inspection of highly variant geared motors,” *Applied Acoustics*, vol. 235, no. 110687, pp. 1–17, 2025
- 6) Y. Sun, Y. Zhou, X. Xu, J. Qi, F. Xu, Z. Ren, and B. W. Schuller, “Weakly-supervised depression detection in speech through self-learning based label correction,” *IEEE Transactions on Audio, Speech and Language Processing*, 2025. 11 pages
- 7) Y. Chang, Z. Ren, Z. Zhang, X. Jing, K. Qian, X. Shao, B. Hu, T. Schultz, and B. W. Schuller, “STAA-Net: A sparse and transferable adversarial attack for speech emotion recognition,” *IEEE Transactions on Affective Computing*, 2024. accepted, to appear
- 8) T. T. Nguyen, T. T. Huynh, Z. Ren, T. T. Nguyen, P. L. Nguyen, H. Yin, and Q. V. H. Nguyen, “Privacy-preserving explainable AI: A survey,” *SCIENCE CHINA Information Sciences*, vol. 68, no. 111101, pp. 1–34, 2025
- 9) P. Wißbrock, Z. Ren, and D. Pelkmann, “More than spectrograms: Deep representation learning for machinery fault detection,” *Applied Acoustics*, vol. 225, no. 110178, pp. 1–14, 2024
- 10) Z. Ren, Y. Chang, T. T. Nguyen, Y. Tan, K. Qian, and B. W. Schuller, “A comprehensive survey on heart sound analysis in the deep learning era,” *IEEE Computational Intelligence Magazine*, vol. 19, no. 3, pp. 42–57, 2024
- 11) T. N. Tam, Z. Ren, T. T. Nguyen, J. Jo, Q. V. H. Nguyen, and H. Yin, “Portable graph-based rumour detection against multimodal heterophily,” *Knowledge-Based Systems*, Dec. 2023
- 12) Z. Ren, Y. Chang, W. Nejdl, and B. Schuller, “Learning complementary representations via attention-based ensemble learning for cough-based COVID-19 recognition,” *Acta Acustica*, vol. 6, no. 29, pp. 1–5, 2022
- 13) Z. Ren, Y. Chang, K. D. Bartl-Pokorny, F. B. Pokorny, and B. Schuller, “The acoustic dissection of cough: Diving into machine listening-based COVID-19 analysis and detection,” *Journal of Voice*, 2022. to appear
- 14) Z. Ren, K. Qian, F. Dong, Z. Dai, W. Nejdl, Y. Yamamoto, and B. Schuller, “Deep attention-based neural networks for explainable heart sound classification,” *Machine Learning with Applications*, vol. 9, pp. 1–9, May 2022
- 15) Y. Chang, X. Jing, Z. Ren, and B. W. Schuller, “CovNet: A transfer learning framework for automatic COVID-19 detection from crowd-sourced cough sounds,” *Frontiers in Digital Health*, vol. 3, pp. 1–11, Jan. 2022
- 16) K. Qian, M. Schmitt, H. Zheng, T. Koike, J. Han, J. Liu, W. Ji, J. Duan, M. Song, Z. Yang, Z. Ren, S. Liu, Z. Zhang, Y. Yamamoto, and B. W. Schuller, “Computer audition for fighting the SARS-CoV-2 corona crisis – Introducing the multi-task speech corpus for COVID-19,” *IEEE Internet of Things Journal*, 2021. 12 pages
- 17) B. Sertolli, Z. Ren, B. W. Schuller, and N. Cummins, “Representation transfer learning from deep end-to-end speech recognition networks for the classification of health states from speech,” *Computer Speech & Language*, vol. 68, p. 101204, July 2021
- 18) M. Song, A. Mallol-Ragolta, E. Parada-Cabaleiro, Z. Yang, S. Liu, Z. Ren, Z. Zhao, and B. W. Schuller, “Frustration recognition from speech during game interaction using wide residual networks,” *Virtual Reality & Intelligent Hardware*, vol. 3, pp. 76–86, Feb. 2021
- 19) Z. Ren, Q. Kong, J. Han, M. Plumley, and B. W. Schuller, “CAA-Net: Conditional atrous CNNs with attention for explainable device-robust acoustic scene classification,” *IEEE Transactions on Multimedia*, vol. 23, pp. 4131–4142, Nov. 2020
- 20) K. Qian, F. Dong, Z. Ren, Z. Dai, B. Dong, and B. W. Schuller, “Opportunities and challenges for heart sound recognition: A brief on the Heart Sounds Shenzhen corpus,” *Journal of Fudan University (Natural Science)*, vol. 59, pp. 354–359, June 2020
- 21) F. Dong, K. Qian, Z. Ren, A. Baird, X. Li, Z. Dai, B. Dong, F. Metze, Y. Yamamoto, and B. W. Schuller, “Machine listening for heart status monitoring: Introducing and benchmarking HSS – the heart sounds shenzhen corpus,” *IEEE Journal of Biomedical and Health Informatics*, vol. 24, pp. 2082–2092, Nov. 2019
- 22) J. Han, Z. Zhang, Z. Ren, and B. W. Schuller, “Exploring perception uncertainty for emotion recognition in dyadic conversation and music listening,” *Cognitive Computation*, vol. 13, pp. 231—240, June 2020
- 23) Z. Zhao, Z. Bao, Y. Zhao, Z. Zhang, N. Cummins, Z. Ren, and B. W. Schuller, “Exploring deep spectrum representations via attention-based recurrent and convolutional neural networks for speech emotion recognition,” *IEEE Access*, vol. 7, pp. 97515–97525, July 2019
- 24) J. Han, Z. Zhang, Z. Ren, and B. W. Schuller, “EmoBed: Strengthening monomodal emotion recognition via training with crossmodal emotion embeddings,” *IEEE Transactions on Affect Computing*, July 2019. 12 pages
- 25) Z. Ren, K. Qian, Z. Zhang, V. Pandit, A. Baird, and B. W. Schuller, “Deep scalogram representations for acoustic scene classification,” *IEEE/CAA Journal of Automatica Sinica*, vol. 5, pp. 662–669, May 2018

## IV. CONFERENCE PAPERS

- 26) R. A. Rammohan, Z. Ren, A. Swiderska, D. Küster, and T. Schultz, “Unveiling deep speech embeddings: Acoustic insights into hatespeech detection,” in *ITG*, 2025
- 27) Y. Chang, Z. Ren, Z. Zhao, T. T. Nguyen, K. Qian, T. Schultz, and B. W. Schuller, “Breaking resource barriers in speech emotion recognition via data distillation,” in *Proc. INTERSPEECH*, (Rotterdam, The Netherlands), pp. 141–145, 2025
- 28) K. Scheck, T. Dombeck, Z. Ren, P. Wu, M. Wand, and T. Schultz, “DiffMV-ETS: Diffusion-based multi-voice electromyography-to-speech conversion using speaker-independent speech training targets,” in *Proc. INTERSPEECH*, (Rotterdam, The Netherlands), pp. 5573–5577, 2025
- 29) Z. Ren, R. A. Rammohan, K. Scheck, S. Li, and T. Schultz, “End-to-end acoustic-linguistic emotion and intent recogni-

- tion enhanced by semi-supervised learning,” in *Proc. EMBC*, (Copenhagen, Denmark), 2025. 4 pages
- 30) R. E. Paul, P. Kock, L. Deichsel, Z. Ren, Y. Hartmann, and T. Schultz, “Validating automated assessment tests with depth sensors in older adults,” in *Proc. EMBC*, (Copenhagen, Denmark), 2025. 4 pages
- 31) R. V. Sharan, Z. Ren, and B. W. Schuller, “Fine-tuning a pre-trained CNN combined with self-attention for speech emotion recognition,” in *Proc. EMBC*, (Copenhagen, Denmark), 2025. 4 pages
- 32) C. Tan, S. Li, Y. Cao, Z. Ren, and T. Schultz, “Investigating effective speaker property privacy protection in federated learning for speech emotion recognition,” in *Proc. ACM Multimedia Asia*, (Kuala Lumpur, Malaysia), 2024
- 33) Z. Ren, K. Scheck, Q. Hou, S. v. Gogh, M. Wand, and T. Schultz, “Diff-ETS: Learning a diffusion probabilistic model for electromyography-to-speech conversion,” in *Proc. EMBC*, 2024
- 34) K. Scheck, Z. Ren, T. Dombeck, J. Sonnert, S. v. Gogh, Q. Hou, M. Wand, and T. Schultz, “Cross-speaker training and adaptation for electromyography-to-speech conversion,” in *Proc. EMBC*, 2024
- 35) Q. Hou, S. v. Gogh, K. Scheck, Z. Ren, T. Schultz, M. Wand, and J. Schmidhuber, “emg2vec: Self-supervised pretraining in electromyography-based silent speech interfaces,” in *Proc. EMBC*, 2024
- 36) M. Song, Z. Yang, A. Triantafyllopoulos, Y. Nakamura, Toru Zhang, Z. Ren, H. Takeuchi, A. Kishi, T. Ishizawa, K. Yoshiuchi, H. Zhang, K. Qian, B. Hu, B. W. Schuller, and Y. Yamamoto, “Crossmodal transformer on multi-physical signals for personalised daily mental health prediction,” in *Proc. ICDMW*, 2023
- 37) Z. Ren, K. Qian, T. Schultz, and B. W. Schuller, “An overview of the icassp special session on ai security and privacy in speech and audio processing,” in *ACM Multimedia Asia workshop*, (Tainan, Taiwan), 2023
- 38) Z. Ren, K. Scheck, and T. Schultz, “Self-learning and active-learning for electromyography-to-speech Conversion,” in *Proc. ITG*, (Achen, Germany), pp. 245–249, 2023
- 39) K. Scheck, D. Ivucic, Z. Ren, and T. Schultz, “Stream-ETS: Low-latency end-to-end speech synthesis from electromyography signals,” in *Proc. ITG*, (Achen, Germany), pp. 200–204, 2023
- 40) L. Basso, Z. Ren, and W. Nejdl, “Towards efficient ECG-based atrial fibrillation detection via parameterised hypercomplex neural networks,” in *Proc. EUSIPCO*, (Helsinki, Finland), pp. 1375–1379, 2023
- 41) Z. Ren, T. T. N. Nguyen, M. M. Zahed, and W. Nejdl, “Self-explaining neural networks for respiratory sound classification with scale-free interpretability,” in *Proc. IJCNN*, (Gold Coast, Australia), 2023. to appear
- 42) T. T. Nguyen, Q.-D. Nguyen, Z. Ren, J. Jo, Q. V. H. Nguyen, and T. T. Nguyen, “10X faster subgraph matching: Dual matching networks with interleaved diffusion attention,” in *Proc. IJCNN*, (Gold Coast, Australia), 2023. to appear
- 43) Z. Ren, T. T. Nguyen, Y. Chang, and B. W. Schuller, “Fast yet effective speech emotion recognition with self-distillation,” in *Proc. ICASSP*, (Rhodes, Greece), 2023. to appear
- 44) Y. Chang, Z. Ren, T. T. Nguyen, K. Qian, and B. W. Schuller, “Knowledge transfer for on-device speech emotion recognition with neural structured learning,” in *Proc. ICASSP*, (Rhodes, Greece), 2023. to appear
- 45) P. Wißbrock, Y. Richter, D. Pelkmann, Z. Ren, and G. Palmer, “Cutting through the noise: An empirical comparison of psychoacoustic and envelope-based features for machinery fault detection,” in *Proc. ICASSP*, (Rhodes, Greece), 2023. to appear
- 46) K. N. D. Quach, Z. Ren, K. V. Tran, V. H. Vu, Y. Chun, T. T. Nguyen, and J. Jo, “Short-term traffic speed prediction using hybrid lstm-svr model,” in *Proc. RiTA*, (Gold Coast, Australia), pp. 438–450, 2022
- 47) Y. Chang, Z. Ren, T. T. Nguyen, W. Nejdl, and B. Schuller, “Example-based explanations with adversarial attacks for respiratory sound analysis,” in *Proc. INTERSPEECH*, (Incheon, Korea), pp. 4003–4007, 2022
- 48) Z. Ren, T. T. Nguyen, and W. Nejdl, “Prototype learning for interpretable respiratory sound analysis,” in *Proc. ICASSP*, (Singapore), pp. 9087–9091, 2022
- 49) T. Yan, H. Meng, S. Liu, E. Parada-Cabaleiro, Z. Ren, and B. W. Schuller, “Convolutional transformer With adaptive position embedding For Covid-19 detection from cough sounds,” in *Proc. ICASSP*, (Singapore), pp. 9092–9096, 2022
- 50) Y. Chang, Z. Ren, and B. W. Schuller, “Transformer-based CNNs: Mining temporal context information for multi-sound COVID-19 diagnosis,” in *Proc. EMBC*, (Virtual), pp. 2335–2338, 2021
- 51) F. Hellmann, Z. Ren, E. André, and B. W. Schuller, “Deformable dilated faster R-CNN for universal lesion detection in CT images,” in *Proc. EMBC*, (Virtual), pp. 2896–2902, 2021
- 52) P. Hecker, F. Pokorny, K. Bartl-Pokorny, U. Reichel, Z. Ren, S. Hantke, F. Eyben, D. Schuller, B. Arnrich, and B. W. Schuller, “Speaking Corona? Human and machine recognition of COVID-19 from voice,” in *Proc. INTERSPEECH*, (Brno, Czech Republic), pp. 1029–1033, 2021
- 53) Z. Ren, J. Han, N. Cummins, and B. W. Schuller, “Enhancing transferability of black-box adversarial attacks via lifelong learning for speech emotion recognition models,” in *Proc. INTERSPEECH*, (Shanghai, China), pp. 496–500, 2020
- 54) J. Han, K. Qian, M. Song, Z. Yang, Z. Ren, S. Liu, J. Liu, H. Zheng, W. Ji, T. Koike, X. Li, Z. Zhang, Y. Yamamoto, and B. W. Schuller, “An early study on intelligent analysis of speech under COVID-19: Severity, sleep quality, fatigue, and anxiety,” in *Proc. INTERSPEECH*, (Shanghai, China), pp. 4946–4950, 2020
- 55) M. Albes, Z. Ren, B. W. Schuller, and N. Cummins, “Squeeze for sneeze: Compact neural networks for cold and flu recognition,” in *Proc. INTERSPEECH*, (Shanghai, China), pp. 4546–4550, 2020
- 56) S. Liu, A. Triantafyllopoulos, Z. Ren, and B. W. Schuller, “Towards speech robustness for acoustic scene classification,” in *Proc. INTERSPEECH*, (Shanghai, China), pp. 3087–3091, 2020
- 57) N. Cummins, Y. Pan, Z. Ren, J. Fritsch, V. Nallanthighal, H. Christensen, D. Blackburn, B. W. Schuller, M. Magimai-Doss, H. Strik, and A. Härmä, “A comparison of acoustic and linguistics methodologies for Alzheimer’s dementia recognition,” in *Proc. INTERSPEECH*, (Shanghai, China), pp. 2182–2186, 2020
- 58) Z. Ren, A. Baird, J. Han, Z. Zhang, and B. W. Schuller, “Generating and protecting against adversarial attacks for deep speech-based emotion recognition models,” in *Proc. ICASSP*, (Barcelona, Spain), pp. 7184–7188, 2020
- 59) Z. Yang, K. Qian, Z. Ren, A. Baird, Z. Zhang, and B. W. Schuller, “Learning multi-resolution representations for acoustic scene classification via neural networks,” in *Proc. CSMT*, (Harbin, China), pp. 133–143, 2019
- 60) K. Qian, Z. Ren, F. Dong, W.-H. Lai, B. W. Schuller, and Y. Yoshiharu, “Deep wavelets for heart sound classification,” in *Proc. ISPACS*, (Taipei, Taiwan), 2019. 2 pages
- 61) Z. Ren, J. Han, N. Cummins, Q. Kong, M. Plumley, and B. W. Schuller, “Multi-instance learning for bipolar disorder diagnosis using weakly labelled speech data,” in *Proc. DPH*, (Marseille, France), pp. 79–83, 2019
- 62) K. Qian, H. Kuromiya, Z. Ren, M. Schmitt, Z. Zhang, T. Nakamura, K. Yoshiuchi, B. W. Schuller, and Y. Yamamoto, “Automatic detection of major depressive disorder via a bag-of-behaviour-words approach,” in *Proc. ISICDM*, (Xi’an, China), pp. 71–75, 2019
- 63) F. Ringeval, B. W. Schuller, M. Valstar, N. Cummins, R. Cowie,

- L. Tavabi, M. Schmitt, S. Alisamir, S. Amiriparian, E.-M. Messner, S. Song, S. Liu, Z. Zhao, A. Mallol-Ragolta, Z. Ren, M. Soleymani, and M. Pantic, “AVEC 2019 workshop and challenge: State-of-mind, detecting depression with AI, and cross-cultural affect recognition,” in *Proc. AVEC*, (Nice, France), pp. 3–12, 2019
- 64) Z. Ren, Q. Kong, J. Han, M. Plumbley, and B. Schuller, “Attention-based atrous convolutional neural networks: Visualisation and understanding perspectives of acoustic scenes,” in *Proc. ICASSP*, (Brighton, UK), pp. 56–60, 2019
- 65) J. Han, Z. Zhang, Z. Ren, and B. Schuller, “Implicit fusion by joint audiovisual training for emotion recognition in mono modality,” in *Proc. ICASSP*, (Brighton, UK), pp. 5861–5865, 2019
- 66) Z. Ren, Q. Kong, K. Qian, M. Plumbley, and B. Schuller, “Attention-based convolutional neural networks for acoustic scene classification,” in *Proc. DCASE*, (Surrey, UK), pp. 39–43, 2018
- 67) Z. Ren, N. Cummins, J. Han, S. Schnieder, J. Krajewski, and B. W. Schuller, “Evaluation of the pain level from speech: Introducing a novel pain database and benchmarks,” in *Proc. ITG*, (Oldenburg, Germany), pp. 56–60, 2018
- 68) J. Han, Z. Zhang, M. Schmitt, Z. Ren, F. Ringeval, and B. W. Schuller, “Bags in bag: Generating context-aware bags for tracking emotions from speech,” in *Proc. INTERSPEECH*, (Hyderabad, India), pp. 3082–3086, 2018
- 69) B. W. Schuller, S. Steidl, A. Batliner, P. Marschik, H. Baumeister, F. Dong, S. Hantke, F. Pokorny, E.-M. Rathner, K. Bartl-Pokorny, C. Einspieler, D. Zhang, A. Baird, S. Amiriparian, K. Qian, Z. Ren, M. Schmitt, P. Tzirakis, and S. Zafeiriou, “The INTERSPEECH 2018 computational paralinguistics challenge: Atypical & self-assessed affect, crying & heart beats,” in *Proc. INTERSPEECH*, (Hyderabad, India), pp. 122–126, 2018
- 70) Z. Ren, N. Cummins, V. Pandit, J. Han, K. Qian, and B. W. Schuller, “Learning image-based representations for heart sound classification,” in *Proc. DH*, (Lyon, France), pp. 143–147, 2018
- 71) J. Han, Z. Zhang, Z. Ren, F. Ringeval, and B. W. Schuller, “Towards conditional adversarial training for predicting emotions from speech,” in *Proc. ICASSP*, (Calgary, Canada), pp. 6822–6826, 2018
- 72) Z. Ren, V. Pandit, K. Qian, Z. Yang, Z. Zhang, and B. W. Schuller, “Deep sequential image features on acoustic scene classification,” in *Proc. DCASE*, (Munich, Germany), pp. 113–117, 2017
- 73) K. Qian, Z. Ren, V. Pandit, Z. Yang, Z. Zhang, and B. W. Schuller, “Wavelets revisited for the classification of acoustic scenes,” in *Proc. DCASE*, (Munich, Germany), pp. 108–112, 2017
- 74) Z. Ren, Q. Zhang, H. Zhu, and Q. Wang, “Extending the FOV from disparity and color consistencies in multiview light fields,” in *Proc. ICIP*, (Beijing, China), pp. 1157–1161, 2017
- 75) J. Deng, N. Cummins, J. Han, X. Xu, Z. Ren, V. Pandit, Z. Zhang, and B. W. Schuller, “The University of Passau open emotion recognition system for the multimodal emotion challenge,” in *Proc. CCPR*, (Chengdu, China), pp. 652–666, 2016
- 76) Y. Zhang, F. Weninger, Z. Ren, and B. W. Schuller, “Sincerity and deception in speech: Two sides of the same coin? A transfer- and multi-task learning perspective,” in *Proc. INTERSPEECH*, (San Francisco, CA), pp. 2041–2045, 2016
- 19 recognition in the DiCOVA 2021 challenge,” tech. rep., DiCOVA Challenge, 2021. 3 pages
- 79) Z. Ren, V. Pandit, K. Qian, Z. Yang, Z. Zhang, and B. W. Schuller, “A system for 2017 DCASE challenge using deep sequential image and wavelet features,” tech. rep., DCASE Challenge, Munich, Germany, 2017. 1 page
- 80) S. Amiriparian, N. Cummins, M. Freitag, K. Qian, Z. Ren, V. Pandit, and B. W. Schuller, “The combined Augsburg/Passau/TUM/ICL system for DCASE 2017,” tech. rep., DCASE Challenge, Munich, Germany, 2017. 1 page

## VII. PATENTS

- 81) Q. Wang, Z. Ren, G. Zhou, and W. Zeng, “Light field acquisition device based on micro camera array and data processing method,” 2022. No. CN106027861A
- 82) Q. Wang, Z. Ji, S. Han, C. Zhang, Z. Ren, and G. Zhou, “An equipment and algorithms for light field camera extrinsic parameters calibration,” 2016. No. CN105654484

## V. DISSERTATION

- 77) Z. Ren, *Deep learning techniques for computer audition*. PhD thesis, Faculty of Applied Computer Science, University of Augsburg, 2022

## VI. TECHNICAL REPORTS

- 78) Z. Ren, Y. Chang, and B. W. Schuller, “The EIHW-GLAM deep attentive multi-model fusion system for cough-based COVID-