1. Record Nr. UNISA996546844803316

Titolo	Man-Machine Speech Communication [[electronic resource]] : 17th National Conference, NCMMSC 2022, Hefei, China, December 15–18, 2022, Proceedings / / edited by Ling Zhenhua, Gao Jianqing, Yu Kai, Jia Jia
Pubbl/distr/stampa	Singapore : , : Springer Nature Singapore : , : Imprint : Springer, , 2023
ISBN	981-9924-01-4
Edizione	[1st ed. 2023.]
Descrizione fisica	1 online resource (342 pages)
Collana	Communications in Computer and Information Science, , 1865-0937 ; ; 1765
Disciplina	006.4
Soggetti	Computer vision Natural language processing (Computer science) Signal processing Artificial intelligence User interfaces (Computer systems) Human-computer interaction Computer Vision Natural Language Processing (NLP) Signal, Speech and Image Processing Artificial Intelligence User Interfaces and Human Computer Interaction
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Nota di bibliografia	Includes bibliographical references and index.
Nota di contenuto	MCPN: A Multiple Cross-Perception Network for Real-Time Emotion Recognition in Conversation -- Baby Cry Recognition Based on Acoustic Segment Model -- A Multi-feature Sets Fusion Strategy with Similar Samples Removal for Snore Sound Classification -- Multi-Hypergraph Neural Networks for Emotion Recognition in Multi-Party Conversations -- Using Emoji as an Emotion Modality in Text-Based Depression Detection -- Source-Filter-Based Generative Adversarial Neural Vocoder for High Fidelity Speech Synthesis -- Semantic enhancement framework for robust speech recognition -- Achieving Timestamp Prediction While Recognizing with Non-Autoregressive End-to-End ASR

Model -- Predictive AutoEncoders are Context-Aware Unsupervised Anomalous Sound Detectors -- A pipelined framework with serialized output training for overlapping speech recognition -- Adversarial Training Based on Meta-Learning in Unseen Domains for Speaker Verification -- Multi-Speaker Multi-Style Speech Synthesis with Timbre and Style Disentanglement -- Multiple Confidence Gates for Joint Training of SE and ASR -- Detecting Escalation Level from Speech with Transfer Learning and Acoustic-Linguistic Information Fusion -- Pre-training Techniques For Improving Text-to-Speech Synthesis By Automatic Speech Recognition Based Data Enhancement -- A Time-Frequency Attention Mechanism with Subsidiary Information for Effective Speech Emotion Recognition -- Interplay between prosody and syntax-semantics: Evidence from the prosodic features of Mandarin tag questions -- Improving Fine-grained Emotion Control and Transfer with Gated Emotion Representations in Speech Synthesis -- Violence Detection through Fusing Visual Information to Auditory Scene -- Mongolian Text-to-Speech Challenge under Low-Resource Scenario for NCMMSC2022 -- VC-AUG Voice Conversion based Data Augmentation for Text-Dependent Speaker Verication -- Transformer-based potential emotional relation mining network for emotion recognition in conversation -- FastFoley Non-Autoregressive Foley Sound Generation Based On Visual Semantics -- Structured Hierarchical Dialogue Policy with Graph Neural Networks -- Deep Reinforcement Learning for On-line Dialogue State Tracking -- Dual Learning for Dialogue State Tracking -- Automatic Stress Annotation and Prediction For Expressive Mandarin TTS -- MnTTS2 An Open-Source Multi-Speaker Mongolian Text-to-Speech Synthesis Dataset.

Sommario/riassunto

This book constitutes the refereed proceedings of the 17th National Conference on Man–Machine Speech Communication, NCMMSC 2022, held in China, in December 2022. The 21 full papers and 7 short papers included in this book were carefully reviewed and selected from 108 submissions. They were organized in topical sections as follows: MCPN: A Multiple Cross-Perception Network for Real-Time Emotion Recognition in Conversation.- Baby Cry Recognition Based on Acoustic Segment Model, MnTTS2 An Open-Source Multi-Speaker Mongolian Text-to-Speech Synthesis Dataset.