1.	Record Nr.	UNINA9910838283003321
	Autore	Jia Jia
	Titolo	Man-Machine Speech Communication [[electronic resource]] : 18th National Conference, NCMMSC 2023, Suzhou, China, December 8–10, 2023, Proceedings / / edited by Jia Jia, Zhenhua Ling, Xie Chen, Ya Li, Zixing Zhang
	Pubbl/distr/stampa	Singapore : , : Springer Nature Singapore : , : Imprint : Springer, , 2024
	ISBN	981-9706-01-7
	Edizione	[1st ed. 2024.]
	Descrizione fisica	1 online resource (380 pages)
	Collana	Communications in Computer and Information Science, , 1865-0937 ; ; 2006
	Altri autori (Persone)	LingZhenhua ChenXie LiYa ZhangZixing
	Disciplina	006.37
	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 contenuto	Ultra-Low Complexity Residue Echo and Noise Suppression Based on Recurrent Neural Network -- Semi-End-to-End Nested Named Entity Recognition from Speech -- A Lightweight Music Source Separation Model with Graph Convolution Network -- Joint time-domain and frequency-domain progressive learning for single-channel speech enhancement and recognition -- A Study on Domain Adaptation for

Audio-visual Speech Enhancement -- APNet2: High-quality and High-efficiency Neural Vocoder with Direct Prediction of Amplitude and Phase Spectra -- Within- and Between-Class Sample Interpolation Based Supervised Metric Learning for Speaker Verification -- Joint speech and noise estimation using SNR-adaptive target learning for deep-learning-based speech enhancement -- Data Augmentation By Finite Element Analysis for Enhanced Machine Anomalous Sound Detection -- A Fast Sampling Method in Diffusion-based Dance Generation Models -- End-to-end Streaming Customizable Keyword Spotting based on text-adaptive neural search -- The Production of Successive Addition Boundary Tone in Mandarin Preschoolers -- Emotional Support Dialog System Through Recursive Interactions Among Large Language Models -- Task-Adaptive Generative Adversarial Network based Speech Dereverberation for Robust Speech Recognition -- Real-time Automotive Engine Sound Simulation with Deep Neural Network -- A Framework Combining Separate and Joint Training for Neural Vocoder-Based Monaural Speech Enhancement -- Accent-VITS: accent transfer for end-to-end TTS -- Multi-branch Network with Cross-Domain Feature Fusion for Anomalous Sound Detection -- A Packet Loss Concealment Method Based on the Demucs Network Structure -- Improving Speech Perceptual Quality and Intelligibility through Sub-band Temporal Envelope Characteristics -- Adaptive Deep Graph Convolutional Network For Dialogical Speech Emotion Recognition -- Iterative Noisy-target Approach: Speech Enhancement without Clean Speech -- Joint Training or Not: An Exploration of Pre-trained Speech Models in Audio-Visual Speaker Diarization -- Zero-shot Singing Voice Conversion Method Based on Timbre Space Modeling and Excitation Signal Control -- A Comparative Study of Pre-trained Audio and Speech Models for Heart Sound Detection -- CAM-GUI: A Conversational Assistant on Mobile GUI -- A Pilot Study on the Prosodic Factors Influencing Voice Attractiveness of AI Speech -- The DKU-MSXF Diarization System for the VoxCeleb Speaker Recognition Challenge 2023 -- Chinese EFL Learners' Auditory and Visual Perception of English Statement and Question Intonation: The Effect of Stress -- An Improved System for Partially Fake Audio Detection Using Pre-trained Model -- Leveraging Synthetic Speech for CIF-based Customized Keyword Spotting.

| Sommario/riassunto | This book constitutes the refereed proceedings of the 18th National Conference on Man-Machine Speech Communication, NCMMSC 2023, held in Suzhou, China, during December 8–11, 2023. The 20 full papers and 11 short papers included in this book were carefully reviewed and selected from 117 submissions. They deal with topics such as speech recognition, synthesis, enhancement and coding, audio/music/singing synthesis, avatar, speaker recognition and verification, human–computer dialogue systems, large language models as well as phonetic and linguistic topics such as speech prosody analysis, pathological speech analysis, experimental phonetics, acoustic scene classification. |