

1. Record Nr.	UNISA996565866403316
Autore	Karpov Alexey
Titolo	Speech and Computer [[electronic resource]] : 25th International Conference, SPECOM 2023, Dharwad, India, November 29 – December 2, 2023, Proceedings, Part I // edited by Alexey Karpov, K. Samudravijaya, K. T. Deepak, Rajesh M. Hegde, Shyam S. Agrawal, S. R. Mahadeva Prasanna
Pubbl/distr/stampa	Cham : , : Springer Nature Switzerland : , : Imprint : Springer, , 2023
ISBN	3-031-48309-X
Edizione	[1st ed. 2023.]
Descrizione fisica	1 online resource (657 pages)
Collana	Lecture Notes in Artificial Intelligence, , 2945-9141 ; ; 14338
Altri autori (Persone)	SamudravijayaK DeepakK. T HegdeRajesh M AgrawalShyam S PrasannaS. R. Mahadeva
Disciplina	006.3
Soggetti	Artificial intelligence Computer engineering Computer networks Application software Image processing - Digital techniques Computer vision Artificial Intelligence Computer Engineering and Networks Computer and Information Systems Applications Computer Imaging, Vision, Pattern Recognition and Graphics
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Nota di contenuto	Automatic Speech Recognition -- Extreme Learning Layer: A Boost for Spoken Digit Recognition with Spiking Neural Networks -- EMO-AVSR: Two-Level Approach for Audio-Visual Emotional Speech Recognition -- Significance of Audio Quality in Speech-to-Text Translation Systems -- Everyday Conversations: a Comparative Study of Expert Transcriptions and ASR Outputs at a Lexical Level -- Improving Automatic Speech

Recognition with Dialect-Specific Language Models -- Emotional speech recognition of Holocaust survivors with deep neural network models for Russian language -- Computational Paralinguistics -- Aggregation Strategies of Wav2vec 2.0 Embeddings for Computational Paralinguistic Tasks -- Rhythm Formant Analysis for Automatic Depression Classification -- Determining Alcohol Intoxication Based on Speech and Neural Networks -- Linear Frequency Residual Cepstral Coefficients for Speech Emotion Recognition -- Enhancing Stutter Detection in Speech using Zero Time Windowing Cepstral Coefficients and Phase Information -- Source and System-based Modulation Approach for Fake Speech Detection -- Digital Signal Processing -- Investigation of Different Calibration Methods for Deep Speaker Embedding based Verification Systems -- Learning to Predict Speech Intelligibility from Speech Distortions -- Sparse Representation Frameworks for Acoustic Scene Classification -- Driver Speech Detection in Real Driving Scenario -- Regularization based Incremental Learning in TCNN for Robust Speech Enhancement Targeting Effective Human Machine Interaction -- Candidate Speech Extraction from Multi-Speaker Single-Channel Audio Interviews -- Post-Processing of Translated Speech by Pole Modification and Residual Enhancement to Improve Perceptual Quality -- Region Normalized Capsule Network based Generative Adversarial Network for Non-Parallel Voice Conversion -- Speech Enhancement using LinkNet Architecture -- ATT: Adversarial Trained Transformer for Speech Enhancement -- Human Identification by Dynamics of Changes in Brain Frequencies Using Artificial Neural Networks -- Speech Prosody -- Analysis of Formant Trajectories of a Speech Signal for the Purpose of Forensic Identification of a Foreign Speaker -- Gestures vs. Prosodic Structure in Laboratory Ironic Speech -- Sounds of < sil > ence: Acoustics of Inhalation in Read Speech -- Prolongations as Hesitation Phenomena in Spoken Speech in First and Second Language -- Study of Indian English Pronunciation Variabilities Relative to Received Pronunciation -- Multimodal Collaboration in Expository Discourse: Verbal and Nonverbal Moves Alignment -- Association of Time Domain Features with Oral Cavity Configuration during Vowel Production and its Application in Vowel Recognition -- Prosodic Interaction Models in a Conversation -- Natural Language Processing -- Development and Research of Dialogue Agents with Long-Term Memory and Web Search -- Pre- and Post-Textual Contexts in Assessment of a Message as Offensive or Defensive Aggression Verbalization -- Boosting Rule-based Grapheme-to-Phoneme Conversion with Morphological Segmentation and Syllabification in Bengali -- Revisiting Assessment of Text Complexity: Lexical and Syntactic Parameters Fluctuations -- Analysis of Natural Language Understanding Systems with L2 Learner Specific Synthetic Grammatical Errors based on Parts-of-Speech -- On the Most Frequent Sequences of Words in Russian Spoken Everyday Language (Bigrams and Trigrams): An Experience of Classification -- Child Speech Processing -- Recognition of the Emotional State of Children by Video and Audio Modalities by Indian and Russian Experts -- Effect of Linear Prediction Order to Modify Formant Locations for Children Speech Recognition -- Gammatone-Filterbank based Pitch-Normalized Cepstral Coefficients for Zero-Resource Children's ASR -- System Assisted Vocal Response Analysis and Assessment of Autism in Children: A Machine Learning Based Approach -- Addressing Effects of Formant Dispersion and Pitch Sensitivity for the Development of Children's KWS System -- Development of Children's KWS System Perceptual Experiment and Automatic Recognition by Video, Audio and Text Modalities -- Linear Frequency Residual Features for Infant Cry

Classification -- Speech Processing for Medicine -- Identification of Voice Disorders: A Comparative Study of Machine Learning Algorithms -- Transfer Learning using Whisper for Dysarthric Automatic Speech Recognition -- Significance of Duration Modification in Reducing Listening Effort of Slurred Speech from Patients with Traumatic Brain Injury -- Significance of Duration Modification in Reducing Listening Effort of Slurred Speech from Patients with Traumatic Brain Injury -- Respiratory Sickness Detection from Audio Recordings using CLIP Models -- Investigating the Effect of Data Impurity on the Detection Performances of Mental Disorders through Spoken Dialogues.

Sommario/riassunto

The two-volume proceedings set LNAI 14338 and 14339 constitutes the refereed proceedings of the 25th International Conference on Speech and Computer, SPECOM 2023, held in Dharwad, India, during November 29–December 2, 2023. The 94 papers included in these proceedings were carefully reviewed and selected from 174 submissions. They focus on all aspects of speech science and technology: automatic speech recognition; computational paralinguistics; digital signal processing; speech prosody; natural language processing; child speech processing; speech processing for medicine; industrial speech and language technology; speech technology for under-resourced languages; speech analysis and synthesis; speaker and language identification, verification and diarization.
