25th Annual Conference of the International Speech Communication Association (INTERSPEECH 2024)

Kos, Greece
1-5 September 2024

Volume 1 of 7

ISBN: 979-8-3313-0506-2

Printed from e-media with permission by:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571

Some format issues inherent in the e-media version may also appear in this print version.

International Speech Communication Association
c/o Mme Emmanuelle FOXONET
4 Rue des Fauvettes - Lous Tourils
F-66390 Baixas, France

Phone: 49 228 735 643
Fax: 33 468 385 827

secretariat@isca-speech.org

Additional copies of this publication are available from:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: 845-758-0400
Fax: 845-758-2633
Email: curran@proceedings.com
Web: www.proceedings.com

TABLE OF CONTENTS

VOLUME 1

SPEECH AND AUDIO ANALYSIS AND REPRESENTATIONS

ACOUSTIC EVENT DETECTION AND CLASSIFICATION 2

DETECTION AND CLASSIFICATION OF BIOACOUSTIC SIGNALS

ACOUSTIC ECHO CANCELLATION

SPEECH SYNTHESIS: VOICE CONVERSION 1

NEURAL NETWORK ARCHITECTURES FOR ASR 2

DECODING ALGORITHMS

PRONUNCIATION ASSESSMENT

SPOKEN LANGUAGE PROCESSING

SPOKEN MACHINE TRANSLATION 2

BIOSIGNAL-ENABLED SPOKEN COMMUNICATION

INDIVIDUAL AND SOCIAL FACTORS IN PHONETICS

PARALINGUISTICS

SPEAKER RECOGNITION: ADVERSARIAL AND SPOOFING ATTACKS

AUDIO EVENT DETECTION AND CLASSIFICATION 1

SOURCE SEPARATION 2

NOISE REDUCTION, DEREVERBERATION, AND ECHO CANCELLATION

VOLUME 2

COMPUTATIONALLY-EFFICIENT SPEECH ENHANCEMENT

ZERO-SHOT TTS

NOISE ROBUSTNESS, FAR-FIELD, AND MULTI-TALKER ASR

CONTEXTUAL BIASING AND ADAPTATION

SPOKEN LANGUAGE UNDERSTANDING

SPOKEN MACHINE TRANSLATION 1

HEARING DISORDERS

SPEECH DISORDERS 2

TAUKADIAL CHALLENGE: SPEECH-BASED COGNITIVE ASSESSMENT IN CHINESE AND ENGLISH (SPECIAL SESSION)

SHOW AND TELL 1

KEYNOTE 2

PHONETICS AND PHONOLOGY OF SECOND LANGUAGE ACQUISITION

CORPORA-BASED APPROACHES IN AUTOMATIC EMOTION RECOGNITION

ANALYSIS OF SPEAKERS STATES AND TRAITS

SPOOFING AND DEEPFAKE DETECTION

AUDIO CAPTIONING, TAGGING, AND AUDIO-TEXT RETRIEVAL

GENERATIVE SPEECH ENHANCEMENT

SPEECH SYNTHESIS: EVALUATION

MULTILINGUAL ASR

GENERAL TOPICS IN ASR

SPOKEN LANGUAGE UNDERSTANDING

SPEECH AND MULTIMODAL RESOURCES

PATHOLOGICAL SPEECH ANALYSIS 1

VOLUME 3

SPEECH AND LANGUAGE IN HEALTH: FROM REMOTE MONITORING TO MEDICAL CONVERSATIONS - 1 (SPECIAL SESSION)

SPEECH AND BRAIN

INNOVATIVE METHODS IN PHONETICS AND PHONOLOGY

SPEAKER AND LANGUAGE IDENTIFICATION AND DIARIZATION

AUDIO-TEXT RETRIEVAL

SPEECH ENHANCEMENT

SPEECH CODING

SPEECH SYNTHESIS: EXPRESSIVITY AND EMOTION

SPEECH SYNTHESIS: TOOLS AND DATA

SPEECH SYNTHESIS: SINGING VOICE SYNTHESIS

LLM IN ASR

VISION AND SPEECH

SPOKEN DOCUMENT SUMMARIZATION

SPEECH AND LANGUAGE IN HEALTH: FROM REMOTE MONITORING TO MEDICAL CONVERSATIONS - 2 (SPECIAL SESSIONS)

SHOW AND TELL 2

PROSODY

FOUNDATIONAL MODELS FOR DEEPFAKE AND SPOOFED SPEECH DETECTION

SPEAKER RECOGNITION 1

SOURCE SEPARATION 1

VOLUME 4

SPEECH SYNTHESIS: PROSODY

ACCENTED SPEECH, PROSODIC FEATURES, DIALECT, EMOTION, SOUND CLASSIFICATION

NEURAL NETWORK ADAPTATION

ASR AND LLMS

PATHOLOGICAL SPEECH ANALYSIS 3

SPEECH DISORDERS 3

SPEECH RECOGNITION WITH LARGE PRETRAINED SPEECH MODELS FOR UNDER-REPRESENTED LANGUAGES (SPECIAL SESSION)

DATABASES AND PROGRESS IN METHODOLOGY

ARTICULATION, CONVERGENCE AND PERCEPTION

SPEECH EMOTION RECOGNITION

SELF-SUPERVISED MODELS IN SPEAKER RECOGNITION

SPEECH QUALITY ASSESSMENT

PRIVACY AND SECURITY IN SPEECH COMMUNICATION 1

SPEECH SYNTHESIS: VOICE CONVERSION 2

SPEECH SYNTHESIS: TEXT PROCESSING

TRAINING METHODS, SELF-SUPERVISED LEARNING, ADAPTATION

NOVEL ARCHITECTURES FOR ASR

MULTIMODALITY AND FOUNDATION MODELS

VOLUME 5

SPOKEN DIALOGUE SYSTEMS AND CONVERSATIONAL ANALYSIS 1

SPEECH TECHNOLOGY

PATHOLOGICAL SPEECH ANALYSIS 2

SPEECH SCIENCE, SPEECH TECHNOLOGY, AND GENDER (SPECIAL SESSION)

SPEECH PRODUCTION AND PERCEPTION

PHONETICS AND PHONOLOGY: SEGMENTALS AND SUPRASEGMENTALS

TOPICS IN PARALINGUISTICS

EMOTION RECOGNITION: FAIRNESS, VARIABILITY, UNCERTAINTY

SPEAKER VERIFICATION

SPATIAL AUDIO AND ACOUSTICS

GENERATIVE MODELS FOR SPEECH AND AUDIO

SPEECH AND AUDIO MODELLING

MULTI-CHANNEL SPEECH ENHANCEMENT

SPEECH SYNTHESIS: PARADIGMS AND METHODS 1

ERROR CORRECTION AND RESCORING

SPOKEN LANGUAGE UNDERSTANDING

SPOKEN DIALOGUE SYSTEMS AND CONVERSATIONAL ANALYSIS 2

COMPUTATIONAL MODELS OF HUMAN LANGUAGE ACQUISITION, PERCEPTION, AND PRODUCTION (SPECIAL SESSION)

SHOW AND TELL 3

NEW AVENUES IN EMOTION RECOGNITION

SPEAKER DIARIZATION 2

SPEAKER RECOGNITION 2

SPEECH AND AUDIO ANALYSIS

SPEECH QUALITY AND INTELLIGIBILITY: PREDICTION AND ENHANCEMENT

SPEECH SYNTHESIS: VOCODERS

ASR MODEL TRAINING METHODS

CROSS-LINGUAL AND MULTILINGUAL PROCESSING

SPEECH ASSESSMENT

QUESTION ANSWERING FROM SPEECH AND SPOKEN DIALOGUE SYSTEMS

SPOKEN DIALOGUE SYSTEMS AND CONVERSATIONAL ANALYSIS 3

DYSARTHRIC SPEECH ASSESSMENT

SPOKEN LANGUAGE MODELS FOR UNIVERSAL SPEECH PROCESSING (SPECIAL SESSION)

KEYNOTE 4

L1/L2 ACQUISITION AND CROSS-LINGUISTIC FACTORS

SPEAKER RECOGNITION EVALUATION AND RESOURCES

SPEECH TYPE CLASSIFICATION

TARGET SPEAKER EXTRACTION

SPEECH SYNTHESIS: VOICE CONVERSION 3

SPEECH SYNTHESIS: PARADIGMS AND METHODS 3

PRIVACY AND SECURITY IN SPEECH COMMUNICATION 2

VOLUME 7

STREAMING ASR

COMPUTATIONAL RESOURCE CONSTRAINED ASR

EVALUATION OF SPEECH TECHNOLOGY SYSTEMS

NEURAL NETWORK TRAINING FOR SPEECH RECOGNITION

LEVERAGING LARGE LANGUAGE MODELS AND CONTEXTUAL FEATURES FOR PHONETIC ANALYSIS (SPECIAL SESSION)

RESPONSIBLE SPEECH FOUNDATION MODELS (SPECIAL SESSION)

MULTIMODAL PARALINGUISTICS

AUTOMATIC EMOTION RECOGNITION

SELF AND WEAKLY-LABELLED SPEAKER VERIFICATION

ACOUSTIC EVENT DETECTION, SEGMENTATION AND CLASSIFICATION

SPEECH AND AUDIO MODELLING

FAKE AUDIO DETECTION

DEEP LEARNING-BASED SPEECH ENHANCEMENT: APPROACHES, SCALABILITY, AND EVALUATION

SPEECH SYNTHESIS: OTHER TOPICS 1

SPEECH SYNTHESIS: OTHER TOPICS 2

SELF-SUPERVISED LEARNING FOR ASR

SPOKEN TERM DETECTION AND SPEECH RETRIEVAL

SPEECH DISORDERS 1

CONNECTING SPEECH-SCIENCE AND SPEECH-TECHNOLOGY FOR CHILDREN'S SPEECH (SPECIAL SESSION)

SHOW AND TELL 4