# "SpeD 2023" PROGRAM

## October 25, 2023

*University "Politehnica" of Bucharest*
**Central Library Building, 4th floor**

**9:30 – 10:00**

Welcome Message from the Organizing Committee and Foreword – **Corneliu Burileanu**, *University "Politehnica" of Bucharest*; **Horia-Nicolai Teodorescu**, *member of the Romanian Academy*, *"Gheorghe Asachi" Technical University of Iaşi.*

## Speech Analysis, Representation and Models. Audio Signal Processing and Applications

### Chair: Constantin Paleologu

**10:00 – 10:20**

- Liveness Detection – Automatic Classification of Spontaneous and Pre-recorded Speech for Biometric Applications
    
    **Cristian-Teodor Neamtu**, **Serban Mihalache**, **Dragos Burileanu**
    
    *Speech and Dialogue Research Lab., University "Politehnica" of Bucharest, Romania*

**10:20 – 10:40**

- How F0 Contours Reflect Speech Nuclear Events
    
    **Doina Jitcă**
    
    *Institute of Computer Science, Romanian Academy – Iaşi Branch, Romania*

**10:40 – 11:00**

- Automatic Audio Upmixing Based on Source Separation and Ambient Extraction Algorithms
    
    **Marian Negru**[1,2], **Bogdan Moroşanu**[1], **Ana Neacşu**[1,2], **Dragoş Drăghicescu**[2], **Cristian Negrescu**[1]
    
    [1] *University "Politehnica" of Bucharest, Romania*
    [2] *Machine Development and Engineering of Neural Networks, Romania*

### 11:00 – 11:30  Coffee Break

**11:30 – 11:50**

- Regularized RLS Adaptive Algorithm with Conjugate Gradient Method
    
    **Cristian-Lucian Stanciu**, **Cristian Anghel**, **Camelia Elisei-Iliescu**
    
    *Dept. of Telecomm., University "Politehnica" of Bucharest, Romania*

**11:50 – 12:10**

- Personalized Multi-Track Leveling Algorithm
    
    **Bogdan Moroşanu**, **Marian Negru**, **Ana Neacşu**, **Cristian Negrescu**, **Constantin Paleologu**
    
    *University "Politehnica" of Bucharest, Romania*

**12:10 – 12:30**

- Classification and Identification of Certain Types of Car Accidents Based on Sound Information

  *Tudor-Vasile Șerban-Moga*, *Lacrimioara Grama*, *Corneliu Rusu*

  *Signal Processing Group, Electronics Dept., Technical University of Cluj-Napoca, Romania*

**12:30 – 14:30  Lunch**

# *Special Session:* Research Advances in Child Speech Technologies

## Chair: Peter Corcoran

**14:30 – 15:20**

- *Invited Paper*: Synthetic Speaking Children – Why We Need Them and How to Make Them

  *Muhammad Ali Farooq*, *Dan Bigioi*, *Rishabh Jain*, *Wang Yao*, *Mariam Yiwere*, *Peter Corcoran*

  *C3 Imaging Research Center, University of Galway, Ireland*

  *Speaker*: *Peter Corcoran*

**15:20 – 15:40**

- A Comparative Analysis between Conformer-Transducer, Whisper, and Wav2vec2 for Improving the Child Speech Recognition

  *Andrei Barcovschi*, *Rishabh Jain*, *Peter Corcoran*

  *C3 Imaging Research Center, University of Galway, Ireland*

**15:40 – 16:00**

- Automated Child Voice Generation: Methodology and Implementation

  *Shaimaa Alwaisi*, *Mohammed Salah Al-Radhi*, *Géza Németh*

  *Dept. of Telecomm. and Media Informatics, Budapest Univ. of Techn. and Economics, Hungary*

**16:00 – 16:30  Coffee Break**

**16:30 – 16:50**

- Improved Child Text-to-Speech Synthesis through Fastpitch-based Transfer Learning

  *Rishabh Jain*, *Peter Corcoran*

  *C3 Imaging Research Center, University of Galway, Ireland*

**16:50 – 17:10**

- Comparison of Data Augmentation Techniques on Filipino ASR for Children's Speech

  *Cathlyn Abion[1]*, *Niel Carlo Lumapag[1]*, *John Cairu Ramirez[1]*, *Christchelle Resulto[1]*, *Crisron Rudolf Lucas[2]*

  [1] *Electrical and Electronics Eng., Univ. of the Philippines - Diliman, Quezon City, Philippines*
  [2] *School of Computer Science, University College Dublin, Ireland*

**17:10 – 17:30**

- Data Center Audio/Video Intelligence on Device (DAVID) - An Edge-AI Platform for Smart-Toys

  *Gabriel Costache[1]*, *Francisco Salgado[1]*, *Rishabh Jain[2]*, *Cosmin Rotariu[1]*, *George Sterpu[1]*, *Peter Corcoran[2]*

  [1] *Xperi Corporation, Galway, Ireland*
  [2] *C3 Imaging Research Center, University of Galway, Ireland*

# October 26, 2023

## Multimodal Processing. Speech Analysis for Medical Applications. Speech Emotion Recognition

**Chairs: Horia-Nicolai Teodorescu, Lăcrimioara Grama**

### 9:30 – 10:20
- *Invited Paper*: Cross-modal understanding and generation of multimodal content
   *Nicu Sebe*
   *University of Trento, Italy*

### 10:20 – 10:40
- Fairness Constraints in Speech-to-Text Applications – Clinical Effects of Salivation
   *Mike H.M. Teodorescu[1], Ionut Taraboanta[2], Sorin Andrian[2], Angela C. Ghiorghe[2], Horia-Nicolai Teodorescu[3,4], Mihaela Grigorie[5]*
   *[1] Information School, University of Washington, Seattle, USA*
   *[2] "Grigore T. Popa" University of Medicine and Pharmacy of Iaşi, Romania*
   *[3] Institute of Computer Science of the Romanian Academy, Iaşi, Romania*
   *[4] Romanian Academy, Iaşi, Romania*
   *[5] Independent Consultant in Psychiatry, Bucharest, Romania*

### 10:40 – 11:00
- Dysarthric Speech Recognition using Depthwise Separable Convolutions: Preliminary Study
   *Seyed Reza Shahamiri[1], Krishnendu Mandal[2], Sudeshna Sarkar[2]*
   *[1] Dept. of Electrical, Computer, and Software Eng., University of Auckland, New Zealand*
   *[2] Computer Science and Eng. Dept., Indian Institute of Technology, Kharagpur, India*

### 11:00 – 11:20
- Analysis of Formant Correlations in Emotional Speech – Do Formants Correlate?
   *Horia-Nicolai Teodorescu[1,2,3], Stefan Andrei Gheltu[1]*
   *[1] Technical University of Iaşi, Romania*
   *[2] Institute of Computer Science of the Romanian Academy, Iaşi, Romania*
   *[3] Romanian Academy, Iaşi, Romania*

### 11:20 – 11:40
- Dysarthric Speech Recognition: A Comparative Study
   *Dhvani Shah, Vanshika Lal, Zihan Zhong, Qianli Wang, Seyed Reza Shahamiri*
   *Dept. of Electrical, Computer, and Software Eng., University of Auckland, New Zealand*

### 11:40 – 12:10  Coffee Break

### 12:10 – 12:30
- Investigation of a Specific Effect of Alcohol on Formants
   *Jan Malucha*
   *Department of Radio Electronics, Brno University of Technology, Czech Republic*

**12:30 – 12:50**

- An Analysis of Large Speech Models-based Representations for Speech Emotion Recognition

  *Adrian Bogdan Stânea*, *Vlad Striletchi*, *Cosmin Striletchi*, *Adriana Stan*

  *Communications Dept., Technical University of Cluj-Napoca, Romania*

**12:50 – 13:10**

- Speech Emotion Recognition for Emergency Services

  *Bogdan Marghescu*, *Ștefan-Adrian Toma*, *Luciana Morogan*, *Ion Bica*

  *Military Technical Academy "Ferdinand I", Bucharest, Romania*

## 13:10 – 15:00  Lunch

# Selected Topics in Signal Processing. Deep Learning Algorithms, Architectures and Applications

### Chair: Horia Cucu

**15:00 – 15:20**

- Analysis of Spectral Entropy and Maximum Power of EEG as Authentication Mechanisms

  *Alexandru Vulpe[1]*, *Marius Zamfirache[1]*, *Alexandru Caranica[2]*

  *[1] Telecomm. Dept., University "Politehnica" of Bucharest, Romania*

  *[2] Speech and Dialogue Research Lab., University "Politehnica" of Bucharest, Romania*

**15:20 – 15:40**

- Applying Multi-objective Acquisition Function Ensemble for a Candidate Proposal Algorithm

  *Cristian Manolache[1]*, *Cristina Maria Andronache[1]*, *Alexandru Caranica[1]*, *Horia Cucu[1]*, *Andi Buzo[2]*, *Cristian Vasile Diaconu[3]*, *Georg Pelz[2]*

  *[1] University "Politehnica" of Bucharest, Romania*

  *[2] Infineon Technologies, Munich, Germany*

  *[3] Infineon Technologies, Bucharest, Romania*

**15:40 – 16:00**

- Experimental Analysis of Network Traffic Databases for Anomaly Detection

  *Maria-Mădălina Andronache[1]*, *Alexandru Vulpe[2]*

  *[1] CAMPUS Research Center, University "Politehnica" of Bucharest, Romania*

  *[2] Telecomm. Dept., University "Politehnica" of Bucharest, Romania*

**16:00 – 16:20**

- A Study on Initial Population Sampling for Multi-Objective Optimization based on Differential Evolution and Bayesian Inference

  *Georgian Nicolae[1]*, *Catalin Visan[1]*, *Dan Curavale[1]*, *Mihai Boldeanu[1]*, *Horia Cucu[1]*, *Andi Buzo[2]*, *Georg Pelz[2]*

  *[1] University "Politehnica" of Bucharest, Romania*

  *[2] Infineon Technologies, Munich, Germany*

## 19:30  SpeD 2023 Banquet

# October 27, 2023

*University "Politehnica" of Bucharest*
**Central Library Building, 4th floor**

## Speech and Language Resources.
## Spoken Language Recognition

### Chair: Silviu-Ioan Bejinariu

**9:30 – 10:20**
- *Invited Paper*: Information Technology and Geolinguistics
  **Silviu-Ioan Bejinariu[1,2], Vasile Apopei[1], Manuela Nevaci[2], Florin-Teodor Olariu[3], Nicolae Saramandu[2]**
  *[1] Institute of Computer Science, Romanian Academy – Iaşi Branch, Romania*
  *[2] "Iorgu Iordan – Al. Rosetti" Institute of Linguistics, Romanian Academy, Bucharest, Romania*
  *[3] "A. Philippide" Institute of Romanian Philology, Romanian Academy – Iaşi Branch, Romania*
  *Speaker*: **Silviu-Ioan Bejinariu**

**10:20 – 10:40**
- Evaluating a Fine-Tuned Whisper Model on Underrepresented Romanian Speech
  **Vasile Păiş, Verginica Barbu Mititelu, Radu Ion, Elena Irimia**
  *Research Institute for Artificial Intelligence "Mihai Drăgănescu", Romanian Academy, Bucharest, Romania*

**10:40 – 11:00**
- Uncovering Bias in ASR Systems: Evaluating Wav2vec2 and Whisper for Dutch Speakers
  **Marcio Fuckner[1], Sophie Horsman[1], Pascal Wiggers[1], Iskaj Janssen[2]**
  *[1] Amsterdam University of Applied Sciences, Netherlands*
  *[2] RTL, Netherlands*

## Text-to-Speech Synthesis

### Chair: Adriana Stan

**11:00 – 11:20**
- Advancing Limited Data Text-to-Speech Synthesis: Non-Autoregressive Transformer for High-Quality Parallel Synthesis
  **Mohammed Salah Al-Radhi[1], Omnia Ibrahim[2], Ali Raheem Mandeel[1], Tamás Gábor Csapó[1], Géza Németh[1]**
  *[1] Dept. of Telecomm. and Media Informatics, Budapest Univ. of Techn. and Economics, Hungary*
  *[2] Dept. of Language Science and Technology, Saarland University, Germany*

### 11:20 – 11:50  Coffee Break

**11:50 – 12:10**
- Enhancing End-to-End Speech Synthesis by Modeling Interrogative Sentences with Speaker Adaptation
  **Ali Raheem Mandeel, Mohammed Salah Al-Radhi, Tamás Gábor Csapó**
  *Dept. of Telecomm. and Media Informatics, Budapest Univ. of Techn. and Economics, Hungary*

**12:10 – 12:30**

- Harmonic-plus-Noise Network with Linear Prediction and Perceptual Weighting Filters
for Filipino Speech Synthesis

  *Camille Marie Tatoy[1], Johanna Lindsey Pasco[1], Josaphat Ira Emmanuel Benedicto[1],*
  *Crisron Rudolf Lucas[2]*
  *[1] Electrical and Electronics Eng., Univ. of the Philippines - Diliman, Quezon City, Philippines*
  *[2] School of Computer Science, University College Dublin, Ireland*

**12:30 – 12:50**

- Modeling Irregular Voice in End-to-End Speech Synthesis via Speaker Adaptation

  *Ali Raheem Mandeel, Mohammed Salah Al-Radhi, Tamás Gábor Csapó*
  *Dept. of Telecomm. and Media Informatics, Budapest Univ. of Techn. and Economics, Hungary*

**12:50 – 13:10**

- Nonparallel Expressive TTS for Unseen Target Speaker using Style-Controlled Adaptive Layer
and Optimized Pitch Embedding

  *Mohammed Salah Al-Radhi, Tamás Gábor Csapó, Géza Németh*
  *Dept. of Telecomm. and Media Informatics, Budapest Univ. of Techn. and Economics, Hungary*

**13:10**   Concluding remarks – *Corneliu Burileanu*, *University "Politehnica" of Bucharest*