# TABLE OF CONTENTS

## AUD-1: AUDIO AND SPEECH SOURCE SEPARATION 1: SPEECH SEPARATION

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>RETHINKING THE SEPARATION LAYERS IN SPEECH SEPARATION NETWORKS</td>
<td>1</td>
</tr>
<tr>
<td>Yi Luo, Zhuo Chen, Cong Han, Chenda Li, Tianyan Zhou, Nima Mesgarani</td>
<td></td>
</tr>
<tr>
<td>ON PERMUTATION INVARIANT TRAINING FOR SPEECH SOURCE SEPARATION</td>
<td>6</td>
</tr>
<tr>
<td>Xiaoyu Liu, Jordi Pons</td>
<td></td>
</tr>
<tr>
<td>COUNT AND SEPARATE: INCORPORATING SPEAKER COUNTING FOR CONTINUOUS SPEAKER SEPARATION</td>
<td>11</td>
</tr>
<tr>
<td>Zhong-Qiu Wang, Deliang Wang</td>
<td></td>
</tr>
<tr>
<td>ULTRA-LIGHTWEIGHT SPEECH SEPARATION VIA GROUP COMMUNICATION</td>
<td>16</td>
</tr>
<tr>
<td>Yi Luo, Cong Han, Nima Mesgarani</td>
<td></td>
</tr>
<tr>
<td>ATTENTION IS ALL YOU NEED IN SPEECH SEPARATION</td>
<td>21</td>
</tr>
<tr>
<td>Cem Subakan, Mirco Ravanelli, Samuele Cornell, Mirko Bronzi, Jianyuan Zhong</td>
<td></td>
</tr>
<tr>
<td>MULTICHANNEL OVERLAPPING SPEAKER SEGMENTATION USING MULTIPLE HYPOTHESIS TRACKING OF ACOUSTIC AND SPATIAL FEATURES</td>
<td>26</td>
</tr>
<tr>
<td>Aidan O. T. Hogg, Christine Evers, Patrick A. Naylor</td>
<td></td>
</tr>
</tbody>
</table>

## AUD-2: AUDIO AND SPEECH SOURCE SEPARATION 2: MUSIC AND SINGING VOICE SEPARATION

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>SEMI-SUPERVISED SINGING VOICE SEPARATION WITH NOISY SELF-TRAINING</td>
<td>31</td>
</tr>
<tr>
<td>Zhepei Wang, Ritwik Giri, Umut Isik, Jean-Marc Valin, Arvindh Krishnaswamy</td>
<td></td>
</tr>
<tr>
<td>NEURO-STEERED MUSIC SOURCE SEPARATION WITH EEG-BASED AUDITORY ATTENTION DECODING AND CONTRASTIVE-NMF</td>
<td>36</td>
</tr>
<tr>
<td>Giorgia Cantisani, Slim Essid, Gaël Richard</td>
<td></td>
</tr>
<tr>
<td>COMPLEX RATIO MASKING FOR SINGING VOICE SEPARATION</td>
<td>41</td>
</tr>
<tr>
<td>Yixuan Zhang, Yuzhou Liu, Deliang Wang</td>
<td></td>
</tr>
<tr>
<td>TRANSCRIPTION IS ALL YOU NEED: LEARNING TO SEPARATE MUSICAL MIXTURES WITH SCORE AS SUPERVISION</td>
<td>46</td>
</tr>
<tr>
<td>Yun-Ning Hung, Gordon Wichern, Jonathan Le Roux</td>
<td></td>
</tr>
<tr>
<td>ALL FOR ONE AND ONE FOR ALL: IMPROVING MUSIC SEPARATION BY BRIDGING NETWORKS</td>
<td>51</td>
</tr>
<tr>
<td>Ryosuke Sawata, Stefan Uhlich, Shusuke Takahashi, Yuki Mitsuifu</td>
<td></td>
</tr>
<tr>
<td>AN HRNET-BLSTM MODEL WITH TWO-STAGE TRAINING FOR SINGING MELODY EXTRACTION</td>
<td>56</td>
</tr>
<tr>
<td>Yongwei Gao, Xingjian Du, Bilei Zhu, Xiaoheng Sun, Wei Li, Zejun Ma</td>
<td></td>
</tr>
</tbody>
</table>
AUD-3: MUSIC SIGNAL ANALYSIS, PROCESSING, AND SYNTHESIS 1: DEEP LEARNING

DEEPF0: END-TO-END FUNDAMENTAL FREQUENCY ESTIMATION FOR MUSIC AND SPEECH SIGNALS ................................................................. 61
Satwinder Singh, Ruili Wang, Yuanhang Qiu

DIFFERENTIABLE SIGNAL PROCESSING WITH BLACK-BOX AUDIO EFFECTS .................... 66
Marco A. Martinez Ramirez, Oliver Wang, Paris Smaragdis, Nicholas J. Bryan

AUTOMATIC MULTITRACK MIXING WITH A DIFFERENTIABLE MIXING CONSOLE OF NEURAL AUDIO EFFECTS ............................................. 71
Christian J. Steinmetz, Jordi Pons, Santiago Pascual, Joan Serrà

SEQUENCE-TO-SEQUENCE SINGING VOICE SYNTHESIS WITH PERCEPTUAL ENTROPY LOSS ........................................................................... 76
Jiatong Shi, Shuai Guo, Nan Huo, Yuekai Zhang, Qin Jin

REVERB CONVERSION OF MIXED VOCAL TRACKS USING AN END-TO-END CONVOLUTIONAL DEEP NEURAL NETWORK ........................................... 81
Junghyan Koo, Seungryeol Paik, Kyogu Lee

EXTENDING MUSIC BASED ON EMOTION AND TONALITY VIA GENERATIVE ADVERSARIAL NETWORK ..................................................... 86
Bo-Wei Tseng, Yih-Liang Shen, Tai-Shih Chi

AUD-4: MUSIC SIGNAL ANALYSIS, PROCESSING, AND SYNTHESIS 2: ANALYSIS AND PROCESSING

IMPROVING THE ROBUSTNESS OF RIGHT WHALE DETECTION IN NOISY CONDITIONS USING DENOISING AUTOENCODERS AND AUGMENTED TRAINING ................ 91
W. Vickers, B. Milner, R. Lee

SELF-SUPERVISED VQ-VAE FOR ONE-SHOT MUSIC STYLE TRANSFER .................................. 96
Ondrej Cífka, Alexey Ozerov, Umut Simsekli, Gaël Richard

CAPTURING TEMPORAL DEPENDENCIES THROUGH FUTURE PREDICTION FOR CNN-BASED AUDIO CLASSIFIERS ............................................. 101
Hongwei Song, Jiqing Han, Shiwen Deng, Zhihao Du

SEGMENTAL DTW: A PARALLELIZABLE ALTERNATIVE TO DYNAMIC TIME WARping ................................................................. 106
Tj Tsai

PITCH-TIMBRE DISENTANgLEMENT OF MUSICAL INSTRUMENT SOUNDS BASED ON VAE-BASED METRIC LEARNING ........................................ 111
Keitaro Tanaka, Ryo Nishikimi, Yoshiaki Bando, Kazuyoshi Yoshii, Shigeo Morishima

AUD-5: ACTIVE NOISE CONTROL, ECHO REDUCTION, AND FEEDBACK REDUCTION 1: ECHO CANCELLATION

ASYNCRONOUS ACOUSTIC ECHO CANCELLATION OVER WIRELESS CHANNELS ................. 116
Robert Ayrapetian, Philip Hilmes, Mohamed Mansour, Trausti Kristjansson, Carlo Murgia
COMBINING ADAPTIVE FILTERING AND COMPLEX-VALUED DEEP POSTFILTERING FOR ACOUSTIC ECHO CANCELLATION

Mhd Modar Halimeh, Thomas Haubner, Annika Briegleb, Alexander Schmidt, Walter Kellermann

DEEP RESIDUAL ECHO SUPPRESSION WITH A TUNABLE TRADEOFF BETWEEN SIGNAL DISTORTION AND ECHO SUPPRESSION

Amir Ivry, Israel Cohen, Baruch Berdugo

ROBUST STFT DOMAIN MULTI-CHANNEL ACOUSTIC ECHO CANCELLATION WITH ADAPTIVE DECORRELATION OF THE REFERENCE SIGNALS

Saeed Bagheri, Daniele Giacobello

A METHOD FOR DETERMINING PERIODICALLY TIME-VARYING BIAS AND ITS APPLICATIONS IN ACOUSTIC FEEDBACK CANCELLATION

Meng Guo

WEIGHTED RECURSIVE LEAST SQUARE FILTER AND NEURAL NETWORK BASED RESIDUAL ECHO SUPPRESSION FOR THE AEC-CHALLENGE

Ziteng Wang, Yueyue Na, Zhang Liu, Biao Tian, Qiang Fu

AUD-6: ACTIVE NOISE CONTROL, ECHO REDUCTION, AND FEEDBACK REDUCTION

ICASSP 2021 ACOUSTIC ECHO CANCELLATION CHALLENGE: INTEGRATED ADAPTIVE ECHO CANCELLATION WITH TIME ALIGNMENT AND DEEP LEARNING-BASED RESIDUAL ECHO PLUS NOISE SUPPRESSION

Renhua Peng, Linjuan Cheng, Chengshi Zheng, Xiaodong Li

ICASSP 2021 ACOUSTIC ECHO CANCELLATION CHALLENGE: DATASETS, TESTING FRAMEWORK, AND RESULTS

Kusha Sridhar, Ross Cutler, Ando Saabas, Tanel Parnamaa, Markus Loide, Hannes Gamper, Sebastian Braun, Robert Aichner, Sriram Srinivasan

AEC IN A NETSHELL: ON TARGET AND TOPOLOGY CHOICES FOR FCRN ACOUSTIC ECHO CANCELLATION

Jan Franzen, Ernst Seidel, Tim Fingscheidt

KERNEL-INTERPOLATION-BASED FILTERED-X LEAST MEAN SQUARE FOR SPATIAL ACTIVE NOISE CONTROL IN TIME DOMAIN

Jesper Brunnström, Shoichi Koyama

WAVE-DOMAIN OPTIMIZATION OF SECONDARY SOURCE PLACEMENT FREE FROM INFORMATION OF ERROR SENSOR POSITIONS

Jian Xu, Kean Chen, Yunhe Li

AUD-7: AUDIO AND SPEECH SOURCE SEPARATION

LASAFT: LATENT SOURCE ATTENTIVE FREQUENCY TRANSFORMATION FOR CONDITIONED SOURCE SEPARATION

Woosung Choi, Minseok Kim, Jaehwa Chung, Soonyoung Jung

SURROGATE SOURCE MODEL LEARNING FOR DETERMINED SOURCE SEPARATION

Robin Scheibler, Masahito Togami
AUDITORY FILTERBANKS BENEFIT UNIVERSAL SOUND SOURCE SEPARATION ......................... 181
Han Li, Kean Chen, Bernhard U. Seeber

WHAT’S ALL THE FUSS ABOUT FREE UNIVERSAL SOUND SEPARATION DATA? ............ 186
Scott Wisdom, Hakan Erdogan, Daniel P. W. Ellis, Romain Serizel, Nicolas Turpault,
Eduardo Fonseca, Justin Salamon, Prem Seetharaman, John R. Hershey

SEPNET: A DEEP SEPARATION MATRIX PREDICTION NETWORK FOR
MULTICHANNEL AUDIO SOURCE SEPARATION ........................................................................ 191
Shota Inoue, Hirokazu Kameoka, Li Li, Shoji Makino

CDPAM: CONTRASTIVE LEARNING FOR PERCEPTUAL AUDIO SIMILARITY ................. 196
Pranay Manocha, Zeyu Jin, Richard Zhang, Adam Finkelstein

AUD-8: AUDIO AND SPEECH SOURCE SEPARATION 4: MULTI-CHANNEL SOURCE
SEPARATION

LINEAR MULTICHANNEL BLIND SOURCE SEPARATION BASED ON TIME-FREQUENCY MASK OBTAINED BY HARMONIC/PERCUSSIVE SOUND SEPARATION .......... 201
Soichiro Oyabu, Daichi Kitamura, Kohei Yatabe

MULTICHANNEL-BASED LEARNING FOR AUDIO OBJECT EXTRACTION .................... 206
Daniel Arteaga, Jordi Pons

DBNET: DOA-DRIVEN BEAMFORMING NETWORK FOR END-TO-END REVERBERANT
SOUND SOURCE SEPARATION .................................................................................................... 211
Ali Aroudi, Sebastian Braun

JOINT DEREVERBERATION AND SEPARATION WITH ITERATIVE SOURCE STEERING ... 216
Taishi Nakashima, Robin Scheibler, Masahito Togami, Nobutaka Ono

EXPLOITING NON-NEGATIVE MATRIX FACTORIZATION FOR BINAURAL SOUND
LOCALIZATION IN THE PRESENCE OF DIRECTIONAL INTERFERENCE ......................... 221
Ingvi Örnolfsson, Torsten Dau, Ning Ma, Tobias May

BLIND EXTRACTION OF MOVING AUDIO SOURCE IN A CHALLENGING
ENVIRONMENT SUPPORTED BY SPEAKER IDENTIFICATION VIA X-VECTORS .................. 226
Jiri Malek, Jakub Jansky, Tomas Kounovsky, Zbynek Koldovsky, Jindrich Zdansky

AUD-9: MUSIC INFORMATION RETRIEVAL AND MUSIC LANGUAGE PROCESSING 1:
BEAT AND MELODY

MIND THE BEAT: DETECTING AUDIO ONSETS FROM EEG RECORDINGS OF MUSIC
LISTENING .................................................................................................................................... 231
Ashvala Vinay, Alexander Lerch, Grace Leslie

DON’T LOOK BACK: AN ONLINE BEAT TRACKING METHOD USING RNN AND
ENHANCED PARTICLE FILTERING ......................................................................................... 236
Mojtaba Heydari, Zhiyao Duan

SINGING MELODY EXTRACTION FROM POLYPHONIC MUSIC BASED ON SPECTRAL
CORRELATION MODELING ........................................................................................................ 241
Xingjian Du, Bilei Zhu, Qiuqiang Kong, Ziejun Ma
IMPROVING AUTOMATIC DRUM TRANSCRIPTION USING LARGE-SCALE AUDIO-TO-MIDI ALIGNED DATA .......................................................... 246
I-Chieh Wei, Chih-Wei Wu, Li Su

FREQUENCY-TEMPORAL ATTENTION NETWORK FOR SINGING MELODY EXTRACTION ........................................................................... 251
Shuai Yu, Xiaoheng Sun, Yi Yu, Wei Li

STATISTICAL CORRECTION OF TRANSCRIBED MELODY NOTES BASED ON PROBABILISTIC INTEGRATION OF A MUSIC LANGUAGE MODEL AND A TRANSCRIPTION ERROR MODEL ........................................................................ 256
Yuki Hiramatsu, Go Shibata, Ryo Nishikimi, Eita Nakamura, Kazuyoshi Yoshii

AUD-10: MUSIC INFORMATION RETRIEVAL AND MUSIC LANGUAGE PROCESSING 2: SINGING VOICE

RELIABILITY ASSESSMENT OF SINGING VOICE F0-ESTIMATES USING MULTIPLE ALGORITHMS .................................................................................................................. 261
Sebastian Rosenzweig, Frank Scherbaum, Meinard Müller

END-TO-END LYRICS RECOGNITION WITH VOICE TO SINGING STYLE TRANSFER ................................................ 266
Sakya Basak, Shruti Agarwal, Sriram Ganapathy, Naoya Takahashi

SINGING LANGUAGE IDENTIFICATION USING A DEEP PHONOTACTIC APPROACH ........................................................................... 271
Lenny Renault, Andrea Vaglio, Romain Hennesquin

ON THE PREPARATION AND VALIDATION OF A LARGE-SCALE DATASET OF SINGING TRANSCRIPTION ........................................................................ 276
Jun-You Wang, Jyh-Shing Roger Jang

JOINT MULTI-PITCH DETECTION AND SCORE TRANSCRIPTION FOR POLYPHONIC PIANO MUSIC ............................................................................. 281
Lele Liu, Veronica Morfi, Emmanouil Benetos

KARAOKE KEY RECOMMENDATION VIA PERSONALIZED COMPETENCE-BASED RATING PREDICTION ........................................................................ 286
Yuan Wang, Shigeki Tanaka, Keita Yokoyama, Hsin-Tai Wu, Yi Fang

AUD-11: AUDITORY MODELING AND HEARING INSTRUMENTS

A CLOSED-LOOP GAIN-CONTROL FEEDBACK MODEL FOR THE MEDIAL EFFERENT SYSTEM OF THE DESCENDING AUDITORY PATHWAY ........................................................................................................ 291
Afagh Farhadi, Skyler G. Jennings, Elizabeth A. Strickland, Laurel H. Carney

DHASP: DIFFERENTIABLE HEARING AID SPEECH PROCESSING ................................................................................................. 296
Zehai Tu, Ning Ma, Jon Barker

COMPUTATIONALLY EFFICIENT DNN-BASED APPROXIMATION OF AN AUDITORY MODEL FOR APPLICATIONS IN SPEECH PROCESSING ........................................................................ 301
Anil Nagathil, Florian Göbel, Alexandru Neluș, Ian C. Bruce

CASCADED ALL-PASS FILTERS WITH RANDOMIZED CENTER FREQUENCIES AND PHASE POLARITY FOR ACOUSTIC AND SPEECH MEASUREMENT AND DATA AUGMENTATION ........................................................................................................ 306
Hideki Kawahara, Kohei Yatabe
PROBING ACOUSTIC REPRESENTATIONS FOR PHONETIC PROPERTIES ................................................................. 311
  Danni Ma, Neville Ryan, Mark Liberman

AN END-TO-END NON-INTRUSIVE MODEL FOR SUBJECTIVE AND OBJECTIVE REAL-WORLD SPEECH ASSESSMENT USING A MULTI-TASK FRAMEWORK ............................................................... 316
  Zhuohuang Zhang, Piyush Vyas, Xuan Dong, Donald S. Williamson

AUD-12: DETECTION AND CLASSIFICATION OF ACOUSTIC SCENES AND EVENTS 1:
FEW SHOT LEARNING

FEW-SHOT CONTINUOUS LEARNING FOR AUDIO CLASSIFICATION ................................................................. 321
  Yu Wang, Nicholas J. Bryan, Mark Cartwright, Juan Pablo Bello, Justin Salamon

ZERO-SHOT AUDIO CLASSIFICATION WITH FACTORED LINEAR AND NONLINEAR
ACOUSTIC-SEMANTIC PROJECTIONS .................................................................................................................. 326
  Huang Xie, Okko Räsänen, Tuomas Virtanen

UNSUPERVISED AND SEMI-SUPERVISED FEW-SHOT ACOUSTIC EVENT
CLASSIFICATION .................................................................................................................................................. 331
  Hsin-Ping Huang, Krishna C. Puvvada, Ming Sun, Chao Wang

FLOW-BASED SELF-SUPERVISED DENSITY ESTIMATION FOR ANOMALOUS SOUND
DETECTION ......................................................................................................................................................... 336
  Kota Dohi, Takashi Endo, Harsh Purohit, Ryo Tanabe, Yohei Kawaguchi

SELF-TRAINING FOR SOUND EVENT DETECTION IN AUDIO MIXTURES .......................................................... 341
  Sangwook Park, Ashwin Bellur, David K. Han, Mounya Elhilali

PROTOTYPICAL NETWORKS FOR DOMAIN ADAPTATION IN ACOUSTIC SCENE
CLASSIFICATION .................................................................................................................................................. 346
  Shubhr Singh, Helen L. Bear, Emmanouil Benetos

AUD-13: DETECTION AND CLASSIFICATION OF ACOUSTIC SCENES AND EVENTS 2:
WEAK SUPERVISION

A GLOBAL-LOCAL ATTENTION FRAMEWORK FOR WEAKLY LABELLED AUDIO
TAGGING ............................................................................................................................................................. 351
  Helin Wang, Yuexian Zou, Wenwu Wang

AN IMPROVED MEAN TEACHER BASED METHOD FOR LARGE SCALE WEAKLY
LABELED SEMI-SUPERVISED SOUND EVENT DETECTION ................................................................................. 356
  Xu Zheng, Yan Song, Ian McLoughlin, Lin Liu, Li-Rong Dai

COMPARISON OF DEEP CO-TRAINING AND MEAN-TEACHER APPROACHES FOR
SEMI-SUPERVISED AUDIO TAGGING ................................................................................................................ 361
  Léo Cances, Thomas Pellegrini

THE BENEFIT OF TEMPORALLY-STRONG LABELS IN AUDIO EVENT CLASSIFICATION ......... 366
  Shawn Hershey, Daniel P W Ellis, Eduardo Fonseca, Aren Jansen, Caroline Liu, R Channing
  Moore, Manoj Plakal

UNSUPERVISED CONTRASTIVE LEARNING OF SOUND EVENT REPRESENTATIONS .............. 371
  Eduardo Fonseca, Diego Ortego, Kevin McGuinness, Noel E. O’Connor, Xavier Serra
<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOUND EVENT DETECTION BY CONSISTENCY TRAINING AND PSEUDO-LABELING...</td>
<td>376</td>
</tr>
<tr>
<td>WITH FEATURE-PYRAMID CONVOLUTIONAL RECURRENT NEURAL NETWORKS</td>
<td></td>
</tr>
<tr>
<td>Chih-Yuan Koh, You-Siang Chen, Yi-Wen Liu, Mingsian R. Bai</td>
<td></td>
</tr>
</tbody>
</table>

**AUD-14: QUALITY AND INTELLIGIBILITY MEASURES**

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>SESQA: SEMI-SUPERVISED LEARNING FOR SPEECH QUALITY ASSESSMENT</td>
<td>381</td>
</tr>
<tr>
<td>Joan Serrà, Jordi Pons, Santiago Pascual</td>
<td></td>
</tr>
<tr>
<td>DETECTING SIGNAL CORRUPTIONS IN VOICE RECORDINGS FOR SPEECH THERAPY</td>
<td>386</td>
</tr>
<tr>
<td>Helmer Nylén, Saikat Chatterjee, Sten Ternström</td>
<td></td>
</tr>
<tr>
<td>MBNET: MOS PREDICTION FOR SYNTHESIZED SPEECH WITH MEAN-BIAS NETWORK</td>
<td>391</td>
</tr>
<tr>
<td>Yichong Leng, Xu Tan, Sheng Zhao, Frank Soong, Xiang-Yang Li, Tao Qin</td>
<td></td>
</tr>
<tr>
<td>NON-INTRUSIVE BINAURAL PREDICTION OF SPEECH INTELLIGIBILITY BASED ON</td>
<td>396</td>
</tr>
<tr>
<td>PHONEME CLASSIFICATION</td>
<td></td>
</tr>
<tr>
<td>Jana Roßbach, Saskia Röttges, Christopher F. Hauth, Thomas Brand, Bernd T. Meyer</td>
<td></td>
</tr>
<tr>
<td>WARP-Q: QUALITY PREDICTION FOR GENERATIVE NEURAL SPEECH CODECS</td>
<td>401</td>
</tr>
<tr>
<td>Wissam A. Jassim, Jan Skoglund, Michael Chinen, Andrew Hines</td>
<td></td>
</tr>
<tr>
<td>CROWDSOURCING APPROACH FOR SUBJECTIVE EVALUATION OF ECHO IMPAIRMENT</td>
<td>406</td>
</tr>
<tr>
<td>Ross Cutler, Babak Nadari, Markus Loide, Sten Sooita, Ando Saabas</td>
<td></td>
</tr>
</tbody>
</table>

**AUD-15: MODELING, ANALYSIS AND SYNTHESIS OF ACOUSTIC ENVIRONMENTS 1:**

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMPLITUDE MATCHING: MAJORIZATION–MINIMIZATION ALGORITHM FOR SOUND FIELD CONTROL ONLY WITH AMPLITUDE CONSTRAINT</td>
<td>411</td>
</tr>
<tr>
<td>Shoichi Koyama, Takashi Amakasu, Natsuki Ueno, Hiroshi Saruwatari</td>
<td></td>
</tr>
<tr>
<td>3D MULTIZONE SOUNDFIELD REPRODUCTION IN A REVERBERANT ENVIRONMENT USING INTENSITY MATCHING METHOD</td>
<td>416</td>
</tr>
<tr>
<td>Huanyu Zuo, Thushara D. Abhayapala, Prasanga N. Samarasinghe</td>
<td></td>
</tr>
<tr>
<td>THE FAR-FIELD EQUATORIAL ARRAY FOR BINAURAL RENDERING</td>
<td>421</td>
</tr>
<tr>
<td>Jens Ahrens, Hannes Helmholz, David Lou Alon, Sebastià V Amengual Gari</td>
<td></td>
</tr>
<tr>
<td>SPHERICAL HARMONIC REPRESENTATION FOR DYNAMIC SOUND-FIELD MEASUREMENTS</td>
<td>426</td>
</tr>
<tr>
<td>Fabrice Katzberg, Marco Maass, Alfred Mertins</td>
<td></td>
</tr>
<tr>
<td>DIRECTION PRESERVING WIND NOISE REDUCTION OF B-FORMAT SIGNALS</td>
<td>431</td>
</tr>
<tr>
<td>Adrian Herzog, Daniele Mirabilii, Emanuël A. P. Habets</td>
<td></td>
</tr>
</tbody>
</table>

**AUD-16: MODELING, ANALYSIS AND SYNTHESIS OF ACOUSTIC ENVIRONMENTS 2:**

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>REFINEMENT OF DIRECTION OF ARRIVAL ESTIMATORS BY MAJORIZATION-MINIMIZATION OPTIMIZATION ON THE ARRAY MANIFOLD</td>
<td>436</td>
</tr>
<tr>
<td>Robin Scheibler, Masahito Togami</td>
<td></td>
</tr>
</tbody>
</table>
ON THE PREDICTABILITY OF HRTFS FROM EAR SHAPES USING DEEP NETWORKS
Yaxuan Zhou, Hao Jiang, Vamsi Krishna Ithapu

APPLIED METHODS FOR SPARSE SAMPLING OF HEAD-RELATED TRANSFER
FUNCTIONS
Lior Arbel, Zamir Ben-Hur, David Lou Alon, Boaz Rafaely

PERSONALIZED HRTF MODELING USING DNN-AUGMENTED BEM
Mengfan Zhang, Jui-Hsien Wang, Doug L. James

EFFICIENT TRAINING DATA GENERATION FOR PHASE-BASED DOA ESTIMATION
Fabian Hübner, Wolfgang Mack, Emanuël A. P. Habets

AUD-17: MODELING, ANALYSIS AND SYNTHESIS OF ACOUSTIC ENVIRONMENTS 3:
ACOUSTIC ANALYSIS

ACOUSTIC REFLECTORS LOCALIZATION FROM STEREO RECORDINGS USING
NEURAL NETWORKS
Giovanni Bologni, Richard Heusdens, Jorge Martinez

DETECTING ACOUSTIC REFLECTORS USING A ROBOT’S EGO-NOISE
Usama Saqib, Antoine Deleforge, Jesper Rindom Jensen

PREDICTION OF OBJECT GEOMETRY FROM ACOUSTIC SCATTERING USING
CONVOLUTIONAL NEURAL NETWORKS
Ziqi Fan, Vibhav Vineet, Chenshen Lu, T. W. Wu, Kyla McMullen

BLIND AMPLITUDE ESTIMATION OF EARLY ROOM REFLECTIONS USING
ALTERNATING LEAST SQUARES
Tom Shlomo, Boaz Rafaely

ACOUSTIC ANALYSIS AND DATASET OF TRANSITIONS BETWEEN COUPLED ROOMS
Thomas McKenzie, Sebastian J. Schlecht, Ville Pulkki

ON LOSS FUNCTIONS FOR DEEP-LEARNING BASED T60 ESTIMATION
Yuying Li, Yuchen Liu, Donald S. Williamson

AUD-18: AUDIO AND SPEECH SOURCE SEPARATION 5: SOURCE SEPARATION

TOWARDS LISTENING TO 10 PEOPLE SIMULTANEOUSLY: AN EFFICIENT
PERMUTATION INVARIANT TRAINING OF AUDIO SOURCE SEPARATION USING
SINKHORN’S ALGORITHM
Hideyuki Tachibana

ACCELERATING AUXILIARY FUNCTION-BASED INDEPENDENT VECTOR ANALYSIS
Andreas Brendel, Walter Kellermann

ONE-SHOT CONDITIONAL AUDIO FILTERING OF ARBITRARY SOUNDS
Beat Gfeller, Dominik Roblek, Marco Tagliasacchi

LOW LATENCY ONLINE BLIND SOURCE SEPARATION BASED ON JOINT
OPTIMIZATION WITH BLIND DEREVERBERATION
Tetsuya Ueda, Tomohiro Nakatani, Rintaro Ikeshita, Keisuke Kinoshita, Shoko Araki, Shoji Makino
AUTOREGRESSIVE FAST MULTICHANNEL NONNEGATIVE MATRIX FACTORIZATION FOR JOINT BLIND SOURCE SEPARATION AND DEREVERBERATION .............................................. 511
Kouhei Sekiguchi, Yoshiaki Bando, Aditya Arie Nugraha, Mathieu Fontaine, Kazuyoshi Yoshii

AUD-19: AUDIO AND SPEECH SOURCE SEPARATION 6: TOPICS IN SOURCE SEPARATION

PHASE RECOVERY WITH BREGMAN DIVERGENCES FOR AUDIO SOURCE SEPARATION .................................................................................................................................................  516
Paul Magron, Pierre-Hugo Vial, Thomas Oberlin, Cédric Févotte

ADVERSARIAL ATTACKS ON AUDIO SOURCE SEPARATION ..........................................................................................................................  521
Naoya Takahashi, Shota Inoue, Yuki Mitsufuji

MAXIMUM A POSTERIORI ESTIMATOR FOR CONVOLUTIVE SOUND SOURCE SEPARATION WITH SUB-SOURCE BASED NTF MODEL AND THE LOCALIZATION PROBABILISTIC PRIOR ON THE MIXING MATRIX ................................................................................ 526
Mieszko Fras, Konrad Kowalczyk

UNIFIED GRADIENT REWEIGHTING FOR MODEL BIASING WITH APPLICATIONS TO SOURCE SEPARATION ................................................................................................................................ 531
Efthymios Tzinis, Dimitrios Bralios, Paris Smaragdis

AUD-20: MUSIC INFORMATION RETRIEVAL AND MUSIC LANGUAGE PROCESSING 3: TOPICS IN MUSIC INFORMATION RETRIEVAL

MELON PLAYLIST DATASET: A PUBLIC DATASET FOR AUDIO-BASED PLAYLIST GENERATION AND MUSIC TAGGING ..........................................................................................................................  536
Andres Ferraro, Yuntae Kim, Soohyeon Lee, Biho Kim, Namjun Jo, Semi Lim, Suyon Lim, Jungtaek Jang, Sehwan Kim, Xavier Serra, Dmitry Bogdanov

INVESTIGATING THE EFFICACY OF MUSIC VERSION RETRIEVAL SYSTEMS FOR SETLIST IDENTIFICATION ..........................................................................................................................  541
Furkan Yesiler, Emilio Molina, Joan Serrà, Emilia Gómez

INSTRUMENT CLASSIFICATION OF SOLO SHEET MUSIC IMAGES ..........................................................................................................................  546
Kevin Ji, Daniel Yang, Tj Tsai

BYTECOVER: COVER SONG IDENTIFICATION VIA MULTI-LOSS TRAINING ..........................................................................................................................  551
Xingjian Du, Zhesong Yu, Bilei Zhu, Xiaou Chen, Zejun Ma

MULTI-TASK SELF-SUPERVISED PRE-TRAINING FOR MUSIC CLASSIFICATION ..........................................................................................................................  556
Ho-Hsiang Wu, Chieh-Chi Kao, Qingming Tang, Ming Sun, Brian McFee, Juan Pablo Bello, Chao Wang

TOWARDS EXPLAINING EXPRESSIVE QUALITIES IN PIANO RECORDINGS: TRANSFER OF EXPLANATORY FEATURES VIA ACOUSTIC DOMAIN ADAPTATION ..........................................................................................................................  561
Shreyan Chowdhury, Gerhard Widmer
AUD-21: MUSIC INFORMATION RETRIEVAL AND MUSIC LANGUAGE PROCESSING 4: STRUCTURE AND ALIGNMENT

SUPERVISED CHORUS DETECTION FOR POPULAR MUSIC USING CONVOLUTIONAL NEURAL NETWORK AND MULTI-TASK LEARNING .................................................................................. 566
Ju-Chiang Wang, Jordan B. L. Smith, Jitong Chen, Xuchen Song, Yuxuan Wang

STRUCTURE-AWARE AUDIO-TO-SCORE ALIGNMENT USING PROGRESSIVELY DILATED CONVOLUTIONAL NEURAL NETWORKS ............................................................................... 571
Ruchit Agrawal, Daniel Wolff, Simon Dixon

LANGUAGE-SENSITIVE MUSIC EMOTION RECOGNITION MODELS: ARE WE REALLY THERE YET? ................................................................................................................................................... 576
Juan Sebastián Gómez-Cañón, Estefanía Cano, Ana Gabriela Pandrea, Perfecto Herrera, Emilia Gómez

LEVERAGING THE STRUCTURE OF MUSICAL PREFERENCE IN CONTENT-AWARE MUSIC RECOMMENDATION ................................................................................................................ 581
Paul Magron, Cédric Févotte

LOW RESOURCE AUDIO-TO-LYRICS ALIGNMENT FROM POLYPHONIC MUSIC RECORDINGS ................................................................................................................................................ 586
Emir Demirel, Sven Ahlbäck, Simon Dixon

MULTIMODAL METRIC LEARNING FOR TAG-BASED MUSIC RETRIEVAL .................................................................................................................. 591
Minz Won, Sergio Oramas, Oriol Nieto, Fabien Gouyon, Xavier Serra


LEARNING CONTEXTUAL TAG EMBEDDINGS FOR CROSS-MODAL ALIGNMENT OF AUDIO AND TAGS ........................................................................................................................................ 596
Xavier Favory, Konstantinos Drossos, Tuomas Virtanen, Xavier Serra

EFFICIENT END-TO-END AUDIO EMBEDDINGS GENERATION FOR AUDIO CLASSIFICATION ON TARGET APPLICATIONS .................................................................................. 601
Paulo Lopez-Meyer, Juan A. Del Hoyo Ontiveros, Hong Lu, Georg Stemmer

TEXT-TO-AUDIO GROUNDING: BUILDING CORRESPONDENCE BETWEEN CAPTIONS AND SOUND EVENTS .................................................................................................................. 606
Xuenan Xu, Heinrich Dinkel, Mengyue Wu, Kai Yu

MULTI-VIEW AUDIO AND MUSIC CLASSIFICATION .......................................................................................................................... 611
Huy Phan, Huy Le Nguyen, Oliver Y. Chén, Lam Pham, Philipp Koch, Ian McLoughlin, Alfred Mertins

AUDIO-VISUAL EVENT RECOGNITION THROUGH THE LENS OF ADVERSARY .......................................................................................................................... 616
Juncheng B Li, Kaixin Ma, Shuhui Qu, Po-Yao Huang, Florian Metze

DCASENET: AN INTEGRATED PRETRAINED DEEP NEURAL NETWORK FOR DETECTING AND CLASSIFYING ACOUSTIC SCENES AND EVENTS .................................................................. 621
Jee-Weon Jung, Hye-Jin Shim, Ju-Ho Kim, Ha-Jin Yu
AUD-23: DETECTION AND CLASSIFICATION OF ACOUSTIC SCENES AND EVENTS 4:
DATASETS AND METRICS

A CURATED DATASET OF URBAN SCENES FOR AUDIO-VISUAL SCENE ANALYSIS .................. 626
Shanshan Wang, Annamaria Mesaros, Toni Heittola, Tuomas Virtanen

IMPROVING SOUND EVENT DETECTION METRICS: INSIGHTS FROM DCASE 2020 ............ 631
Giacomo Ferroni, Nicolas Turpault, Juan Azcarreta, Francesco Tuveri, Romain Serizel,
Çagdas Bilen, Sacha Krstulovic

ARTIFICIALLY SYNTHESISING DATA FOR AUDIO CLASSIFICATION AND
SEGMENTATION TO IMPROVE SPEECH AND MUSIC DETECTION IN RADIO
BROADCAST ......................................................................................................................... 636
Satvik Venkatesh, David Moffat, Alexis Kirke, Gözel Shakeri, Stephen Brewster, Jörg
Fachner, Helen Odell-Miller, Alex Street, Nicolas Farina, Sube Banerjee, Eduardo Reck
Miranda

LSSED: A LARGE-SCALE DATASET AND BENCHMARK FOR SPEECH EMOTION
RECOGNITION ..................................................................................................................... 641
Weiquan Fan, Xiangmin Xu, Xiaofen Xing, Weidong Chen, Dongyan Huang

ENHANCING AUDIO AUGMENTATION METHODS WITH CONSISTENCY LEARNING .......... 646
Turab Iqbal, Karim Helwani, Arvindh Krishnaswamy, Wenwu Wang

FAST THRESHOLD OPTIMIZATION FOR MULTI-LABEL AUDIO TAGGING USING
SURROGATE GRADIENT LEARNING ............................................................................. 651
Thomas Pellegrini, Timothée Masquelier

AUD-24: SIGNAL ENHANCEMENT AND RESTORATION 1: DEEP LEARNING

TOWARDS EFFICIENT MODELS FOR REAL-TIME DEEP NOISE SUPPRESSION ............... 656
Sebastian Braun, Hannes Gamper, Chandan K. A. Reddy, Ivan Tashev

TEACHER-Student LEARNING FOR LOW-LATENCY ONLINE SPEECH
ENHANCEMENT USING WAVE-U-NET ........................................................................... 661
Sotaro Nakaoka, Li Li, Shota Inoue, Shoji Makino

LEARNING DISENTANGLLED FEATURE REPRESENTATIONS FOR SPEECH
ENHANCEMENT VIA ADVERSARIAL TRAINING .......................................................... 666
Nana Hou, Chenglin Xu, Eng Siong Chng, Haizhou Li

SPEECH ENHANCEMENT AUTOENCODER WITH HIERARCHICAL LATENT
STRUCTURE ....................................................................................................................... 671
Koen Oostermeijer, Jun Du, Qing Wang, Chin-Hui Lee

VARIATIONAL AUTOENCODER FOR SPEECH ENHANCEMENT WITH A NOISE-AWARE
ENCODER ......................................................................................................................... 676
Huajian Fang, Guillaume Carbajal, Stefan Wermter, Timo Gerkmann

GUIDED VARIATIONAL AUTOENCODER FOR SPEECH ENHANCEMENT WITH A
SUPERVISED CLASSIFIER ........................................................................................... 681
Guillaume Carbajal, Julius Richter, Timo Gerkmann
### AUD-25: SIGNAL ENHANCEMENT AND RESTORATION 2: AUDIO CODING AND RESTORATION

**An Extension of Sparse Audio Declipper to Multiple Measurement Vectors**

Satoru Emura, Noboru Harada  .................................................................................................................. 686

**Real-Time Speech Frequency Bandwidth Extension**

Yunpeng Li, Marco Tagliasacchi, Oleg Rybakov, Victor Ungureanu, Dominik Roblek  ................................................................. 691

**Bandwidth Extension Is All You Need**

Jiaqi Su, Yunyun Wang, Adam Finkelstein, Zeyu Jin  ..................................................................................... 696

**Audio Dequantization Using (Co)Sparse (Non)Convex Methods**

Pavel Záviška, Pavel Rajmic, Ondrej Mokrý  .................................................................................................................. 701

**Source-Aware Neural Speech Coding for Noisy Speech Compression**

Haici Yang, Kai Zhen, Seungkwan Beach, Minje Kim  .......................................................................................... 706

**Enhancing Into the Codec: Noise Robust Speech Coding with Vector-Quantized Autoencoders**

Jonah Casebeer, Vinjai Vale, Umut Isik, Jean-Marc Valin, Ritwik Giri, Arvindh Krishnaswamy  .......................................................... 711

### AUD-26: SIGNAL ENHANCEMENT AND RESTORATION 3: SIGNAL ENHANCEMENT

**Speech Enhancement with Mixture of Deep Experts with Clean Clustering Pre-Training**

Shlomo E. Chazan, Jacob Goldberger, Sharon Gannot  ..................................................................................... 716

**A Novel NMF-HMM Speech Enhancement Algorithm Based on Poisson Mixture Model**

Yang Xiang, Liming Shi, Jesper Lisby Højvang, Morten Højfeldt Rasmussen, Mads Græsbøll Christensen  ................................................................. 721

**Phoneme-Based Distribution Regularization for Speech Enhancement**

Yajing Liu, Xiulian Peng, Zhiewei Xiong, Yan Lu  .......................................................................................... 726

**Compressed Representation of Cepstral Coefficients via Recurrent Neural Networks for Informed Speech Enhancement**

Carol Chermaz, Dario Leuchtmann, Simon Tanner, Roger Wattenhofer  ................................................................. 731

**Optimizing Short-Time Fourier Transform Parameters via Gradient Descent**

An Zhao, Krishna Subramani, Paris Smaragdis  .......................................................................................... 736

### AUD-27: ACOUSTIC SENSOR ARRAY PROCESSING 1: ARRAY DESIGN AND CALIBRATION

**Iterative Geometry Calibration from Distance Estimates for Wireless Acoustic Sensor Networks**

Tobias Gburrek, Joerg Schmalenstroer, Reinhold Haeb-Umbach  ..................................................................................... 741
ON THE DESIGN OF SQUARE DIFFERENTIAL MICROPHONE ARRAYS WITH A MULTISTAGE STRUCTURE ........................................................................................................................  746
    Xudong Zhao, Gongping Huang, Jacob Benesty, Jingdong Chen, Israel Cohen

ARRAYS OF FIRST-ORDER STEERABLE DIFFERENTIAL MICROPHONES .......................................................... 751
    Federico Borra, Alberto Bernardini, Ivan Bertuletti, Fabio Antonacci, Augusto Sarti

PLANAR ARRAY GEOMETRY OPTIMIZATION FOR REGION SOUND ACQUISITION........................................... 756
    Xi Chen, Chao Pan, Jingdong Chen, Jacob Benesty

ESTIMATION OF MICROPHONE CLUSTERS IN ACOUSTIC SENSOR NETWORKS USING UNSUPERVISED FEDERATED LEARNING.................................................................................................. 761
    Alexandru Nelus, Rene Glitza, Rainer Martin

MISALIGNMENT RECOGNITION IN ACOUSTIC SENSOR NETWORKS USING A SEMI-SUPERVISED SOURCE ESTIMATION METHOD AND MARKOV RANDOM FIELDS...................................................... 766
    Gabriel F Miller, Andreas Brendel, Walter Kellermann, Sharon Gannot

AUD-28: ACOUSTIC SENSOR ARRAY PROCESSING 2: BEAMFORMING

ROTATION-ROBUST BEAMFORMING BASED ON SOUND FIELD INTERPOLATION WITH REGULARLY CIRCULAR MICROPHONE ARRAY ........................................................................................................ 771
    Yukoh Wakabayashi, Kouei Yamaoka, Nobutaka Ono

SPARSE RECOVERY BEAMFORMING AND UPSCALING IN THE RAY SPACE..................................................... 776
    S. Yu, C. Jin, F. Antonacci, A. Sarti

COMBINED DIFFERENTIAL BEAMFORMING WITH UNIFORM LINEAR MICROPHONE ARRAYS.................................................. 781
    Gongping Huang, Yuzhu Wang, Jacob Benesty, Israel Cohen, Jingdong Chen

POLYNOMIAL MATRIX EIGENVALUE DECOMPOSITION OF SPHERICAL HARMONICS FOR SPEECH ENHANCEMENT ................................................................................................................ 786
    Vincent W. Neo, Christine Evers, Patrick A. Naylor

A PARAMETRIC UNCONSTRAINED BINAURAL BEAMFORMER BASED NOISE REDUCTION AND SPATIAL CUE PRESERVATION FOR HEARING-ASSISTIVE DEVICES........................................... 791
    Jie Zhang

A SIMPLIFIED WIENER BEAMFORMER BASED ON COVARIANCE MATRIX MODELLING .................................................. 796
    Fan Zhang, Chao Pan, Jacob Benesty, Jingdong Chen

AUD-29: ACOUSTIC SENSOR ARRAY PROCESSING 3: ACOUSTIC SENSOR ARRAYS

CONTROL ARCHITECTURE OF THE DOUBLE-CROSS-CORRELATION PROCESSOR FOR SAMPLING-RATE-OFFSET ESTIMATION IN ACOUSTIC SENSOR NETWORKS............................................. 801
    Aleksej Chinaev, Sven Wienand, Gerald Enzner

DEFICIENT BASIS ESTIMATION OF NOISE SPATIAL COVARIANCE MATRIX FOR RANK-CONSTRAINED SPATIAL COVARIANCE MATRIX ESTIMATION METHOD IN BLIND SPEECH EXTRACTION.................................................................................. 806
    Yuto Kondo, Yuki Kubo, Norihiro Takamune, Daichi Kitamura, Hiroshi Saruwatari
<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>REDUCING MODAL ERROR PROPAGATION THROUGH CORRECTING MISMATCHED</td>
<td>811</td>
</tr>
<tr>
<td>MICROPHONE GAINS USING RAPID</td>
<td></td>
</tr>
<tr>
<td>Noman Akbar, Glenn Dickins, Mark R. P. Thomas, Prasanga Samarasinghe,</td>
<td></td>
</tr>
<tr>
<td>Thushara Abhayapala</td>
<td></td>
</tr>
<tr>
<td>EVALUATION AND COMPARISON OF THREE SOURCE DIRECTION-OF-ARRIVAL</td>
<td>815</td>
</tr>
<tr>
<td>ESTIMATORS USING RELATIVE HARMONIC COEFFICIENTS</td>
<td></td>
</tr>
<tr>
<td>Yonggang Hu, Prasanga N. Samarasinghe, Sharon Gannot, Thushara D.</td>
<td></td>
</tr>
<tr>
<td>Abhayapala</td>
<td></td>
</tr>
<tr>
<td>NETWORK-AWARE OPTIMAL MICROPHONE CHANNEL SELECTION IN WIRELESS</td>
<td>820</td>
</tr>
<tr>
<td>ACOUSTIC SENSOR NETWORKS</td>
<td></td>
</tr>
<tr>
<td>Michael Gunther, Haitham Afifi, Andreas Brendel, Holger Karl, Walter</td>
<td></td>
</tr>
<tr>
<td>Kellermann</td>
<td></td>
</tr>
<tr>
<td>SUPERVISED DIRECT-PATH RELATIVE TRANSFER FUNCTION LEARNING FOR</td>
<td>825</td>
</tr>
<tr>
<td>BINAURAL SOUND SOURCE LOCALIZATION</td>
<td></td>
</tr>
<tr>
<td>Bing Yang, Xiaofei Li, Hong Liu</td>
<td></td>
</tr>
<tr>
<td>AUD-30: DETECTION AND CLASSIFICATION OF ACOUSTIC SCENES AND EVENTS 5:</td>
<td></td>
</tr>
<tr>
<td>SCENES</td>
<td></td>
</tr>
<tr>
<td>CROSS-MODAL SPECTRUM TRANSFORMATION NETWORK FOR ACOUSTIC SCENE</td>
<td>830</td>
</tr>
<tr>
<td>CLASSIFICATION</td>
<td></td>
</tr>
<tr>
<td>Yang Liu, Alexandras Neophytou, Sunando Sengupta, Eric Sommerlade</td>
<td></td>
</tr>
<tr>
<td>DOMESTIC ACTIVITIES CLUSTERING FROM AUDIO RECORDINGS USING</td>
<td>835</td>
</tr>
<tr>
<td>CONVOLUTIONAL CAPSULE AUTOENCODER NETWORK</td>
<td></td>
</tr>
<tr>
<td>Ziheng Lin, Yaxiong Li, Zhangjin Huang, Wenhao Zhang, Yufeng Tan,</td>
<td></td>
</tr>
<tr>
<td>Yichun Chen, Qianhua He</td>
<td></td>
</tr>
<tr>
<td>SOUND EVENT DETECTION AND SEPARATION: A BENCHMARK ON DESED SYNTHETIC</td>
<td>840</td>
</tr>
<tr>
<td>SOUNDSCAPES</td>
<td></td>
</tr>
<tr>
<td>Nicolas Turpault, Romain Serizel, Scott Wisdom, Hakan Erdogan, John</td>
<td></td>
</tr>
<tr>
<td>R. Hershey, Eduardo Fonseca, Prem Seetharaman, Justin Salamon</td>
<td></td>
</tr>
<tr>
<td>A TWO-STAGE APPROACH TO DEVICE-ROBUST ACOUSTIC SCENE CLASSIFICATION</td>
<td>845</td>
</tr>
<tr>
<td>Hu Hu, Chao-Han Huck Yang, Xianjun Xia, Yue Bai, Xin Tang, Yajian</td>
<td></td>
</tr>
<tr>
<td>Wang, Shutong Niu, Li Chai, Juanjuan Li, Hongning Zhu, Feng Bao,</td>
<td></td>
</tr>
<tr>
<td>Yuanjun Zhao, Sabato Marco Siniscalchi, Yannan Wang, Jun Du, Chin-Hui</td>
<td></td>
</tr>
<tr>
<td>Lee</td>
<td></td>
</tr>
<tr>
<td>SUBSPECTRAL NORMALIZATION FOR NEURAL AUDIO DATA PROCESSING</td>
<td>850</td>
</tr>
<tr>
<td>Simyung Chang, Hyoungwoo Park, Janghoon Cho, Hyunsin Park, Sungrack</td>
<td></td>
</tr>
<tr>
<td>Yun, Kyuwoong Hwang</td>
<td></td>
</tr>
<tr>
<td>SLOW-FAST AUDITORY STREAMS FOR AUDIO RECOGNITION</td>
<td>855</td>
</tr>
<tr>
<td>Evangelos Kazakos, Arsha Nagrani, Andrew Zisserman, Dima Damen</td>
<td></td>
</tr>
<tr>
<td>AUD-31: DETECTION AND CLASSIFICATION OF ACOUSTIC SCENES AND EVENTS 6:</td>
<td></td>
</tr>
<tr>
<td>EVENTS</td>
<td></td>
</tr>
<tr>
<td>IMPACT OF SOUND DURATION AND INACTIVE FRAMES ON SOUND EVENT</td>
<td>860</td>
</tr>
<tr>
<td>DETECTION PERFORMANCE</td>
<td></td>
</tr>
<tr>
<td>Keisuke Imoto, Sakiko Mishima, Yumi Arai, Reishi Kondo</td>
<td></td>
</tr>
</tbody>
</table>
A NEW DCASE 2017 RARE SOUND EVENT DETECTION BENCHMARK UNDER EQUAL TRAINING DATA: CRNN WITH MULTI-WIDTH KERNELS ................................................................. 865
Jan Baumann, Patrick Meyer, Timo Lohrenz, Alexander Roy, Michael Papendieck, Tim Fingscheidt

ROOM ADAPTIVE CONDITIONING METHOD FOR SOUND EVENT CLASSIFICATION IN REVERBERANT ENVIRONMENTS ................................................................. 870
Jaejun Lee, Donmoon Lee, Hyeong-Seok Choi, Kyogu Lee

SOUND EVENT DETECTION BASED ON CURRICULUM LEARNING CONSIDERING LEARNING DIFFICULTY OF EVENTS .................................................................................. 875
Noriyuki Tonami, Keisuke Imoto, Yuki Okamoto, Takahiro Fukumori, Yoichi Yamashita

SOUND EVENT DETECTION IN URBAN AUDIO WITH SINGLE AND MULTI-RATE PCEN.......... 880
Christopher Ick, Brian McFee

AN IMPROVED EVENT-INDEPENDENT NETWORK FOR POLYPHONIC SOUND EVENT LOCALIZATION AND DETECTION ........................................................................ 885
Yin Cao, Turab Iqbal, Qiugiang Kong, Fengyan An, Wenwu Wang, Mark D. Plumbley

AUD-32: AUDIO FOR MULTIMEDIA AND AUDIO PROCESSING SYSTEMS

LIGHTWEIGHT AND INTERPRETABLE NEURAL MODELING OF AN AUDIO DISTORTION EFFECT USING HYPERCONDITIONED DIFFERENTIABLE BIQUADS......................... 890
Shahan Nercessian, Andy Sarroff, Kurt James Werner

ATTACKING AND DEFENDING BEHIND A PSYCHOACOUSTICS-BASED CAPTCHA ................. 895
Chih-Hsiang Huang, Po-Hao Wu, Yi-Wen Liu, Shan-Hung Wu

DOUBLE-DCCCAE: ESTIMATION OF BODY GESTURES FROM SPEECH WAVEFORM ............. 900
Jinhong Lu, Tianhang Liu, Shuzhuang Xu, Hiroshi Shimodaira

INVESTIGATING LOCAL AND GLOBAL INFORMATION FOR AUTOMATED AUDIO CAPTIONING WITH TRANSFER LEARNING ................................................................. 905
Xuenan Xu, Heinrich Dinkel, Mengyue Wu, Zeyu Xie, Kai Yu

AUD-33: TOPICS IN DEEP LEARNING FOR SPEECH AND AUDIO

UNIDIRECTIONAL MEMORY-SELF-ATTENTION TRANSDUCER FOR ONLINE SPEECH RECOGNITION .............................................................................................................. 910
Jian Luo, Jianzong Wang, Ning Cheng, Jing Xiao

ACCDOA: ACTIVITY-COUPLED CARTESIAN DIRECTION OF ARRIVAL REPRESENTATION FOR SOUND EVENT LOCALIZATION AND DETECTION ........................................ 915
Kazuki Shimada, Yuichiro Koyama, Naoya Takahashi, Shusuke Takahashi, Yuki Mitsufuji

SEEN AND UNSEEN EMOTIONAL STYLE TRANSFER FOR VOICE CONVERSION WITH A NEW EMOTIONAL SPEECH DATASET ................................................................. 920
Kun Zhou, Berrak Sisman, Rui Liu, Haizhou Li

U-CONVOLUTION BASED RESIDUAL ECHO SUPPRESSION WITH MULTIPLE ENCODERS ................................................................................................................................. 925
Eesung Kim, Jae-Jin Jeon, Hyeji Seo
A MULTI-CHANNEL TEMPORAL ATTENTION CONVOLUTIONAL NEURAL NETWORK MODEL FOR ENVIRONMENTAL SOUND CLASSIFICATION
You Wang, Chuyao Feng, David V. Anderson

A GENERAL NETWORK ARCHITECTURE FOR SOUND EVENT LOCALIZATION AND DETECTION USING TRANSFER LEARNING AND RECURRENT NEURAL NETWORK
Thi Ngoc Tho Nguyen, Ngoc Khanh Nguyen, Huy Phan, Lam Pham, Kenneth Ooi, Douglas L. Jones, Woon-Seng Gan

AUD-34: ACOUSTIC SYSTEM IDENTIFICATION AND MODELING
ROBUST RECURSIVE LEAST M-ESTIMATE ADAPTIVE FILTER FOR THE IDENTIFICATION OF LOW-RANK ACOUSTIC SYSTEMS
Hongsen He, Jingdong Chen, Jacob Benesty, Yi Yu

NOISE-ROBUST ADAPTATION CONTROL FOR SUPERVISED ACOUSTIC SYSTEM IDENTIFICATION EXPLOITING A NOISE DICTIONARY
Thomas Haubner, Andreas Brendel, Mohamed Elminshawi, Walter Kellermann

INTERPOLATION OF IRREGULARLY SAMPLED FREQUENCY RESPONSE FUNCTIONS USING CONVOLUTIONAL NEURAL NETWORKS
M. Acerbi, R. Malvermi, M. Pezzoli, F. Antonacci, A. Sarti, R. Corradi

EFFECTIVE RANK-BASED ESTIMATION OF THE COHERENT-TO-DIFFUSE POWER RATIO
Heinrich W. Löllmann, Andreas Brendel, Walter Kellermann

ROOM IMPULSE RESPONSE INTERPOLATION FROM A SPARSE SET OF MEASUREMENTS USING A MODAL ARCHITECTURE
Orchisama Das, Paul Calamia, Sebastia V. Amengual Gari

PROCESSING PIPELINES FOR EFFICIENT, PHYSICALLY-ACCURATE SIMULATION OF MICROPHONE ARRAY SIGNALS IN DYNAMIC SOUND SCENES
Alastair H. Moore, Rebecca R. Vos, Patrick A. Naylor, Mike Brookes

BIO-1: BRAIN-COMPUTER INTERFACES
A CLASSIFIER FOR IMPROVING CAUSE AND EFFECT IN SSVEP-BASED BCIS FOR INDIVIDUALS WITH COMPLEX COMMUNICATION DISORDERS
Hadi Habibzadeh, Olivia Zhou, James J. S. Norton, Theresa M. Vaughan, Daphney-Stavroula Zois

SAGA: SPARSE ADVERSARIAL ATTACK ON EEG-BASED BRAIN COMPUTER INTERFACE
Boyuan Feng, Yuke Wang, Yufei Ding

RIEMANNIAN GEOMETRY ON CONNECTIVITY FOR CLINICAL BCI
Marie-Constance Corsi, Florian Yger, Sylvain Chevallier, Camille Noûs

DECODING MUSIC ATTENTION FROM “EEG HEADPHONES”: A USER-FRIENDLY AUDITORY BRAIN-COMPUTER INTERFACE
Winko W. An, Barbara Shinn-Cunningham, Hannes Gamper, Dimitra Emmanouilidou, David Johnston, Mihai Jalobeanu, Edward Cutrell, Andrew Wilson, Kuan-Jung Chiang, Ivan Tashev
MITIGATING INTER-SUBJECT BRAIN SIGNAL VARIABILITY FOR EEG-BASED DRIVER FATIGUE STATE CLASSIFICATION ................................................................. 990
   Sunhee Hwang, Sungho Park, Dohyung Kim, Jewook Lee, Hyeran Byun

A DEEP SPATIO-TEMPORAL MODEL FOR EEG-BASED IMAGINED SPEECH RECOGNITION ............................................................ 995
   Pradeep Kumar, Erik Scheme

**BIO-2: BIOMEDICAL SIGNAL PROCESSING: DETECTION AND ESTIMATION**

INCORPORATING UNCERTAINTY IN DATA LABELING INTO DETECTION OF BRAIN INTERICTAL EPILEPTIFORM DISCHARGES FROM EEG USING WEIGHTED OPTIMIZATION .............................................................................................................. 1000
   Bahman Abdi-Sargezeh, Antonio Valentin, Gonzalo Alarcon, Saeid Sanei

MULTI-LEVEL REVERSIBLE ENCRYPTION FOR ECG SIGNALS USING COMPRESSION SENSING ................................................................. 1005
   Mikko Impiö, Mehmet Yamaç, Jenni Raitoharju

VALIDATING THE INSPIRED SINEWAVE TECHNIQUE TO MEASURE LUNG HETEROGENEITY COMPARED TO ATELECTASIS & OVER-DISTENDED VOLUME IN COMPUTED TOMOGRAPHY IMAGES ........................................................................................................... 1010
   Minh C. Tran, Phi-Anh Phan, Douglas C. Crockett, Federico Formenti, John N. Cronin, Stephen J. Payne, Andrew D. Farmery

A PATIENT-INVARIENT MODEL FOR FREEZING OF GAIT DETECTION AIDED BY WAVELET DECOMPOSITION ................................................................. 1015
   Nasimuddin Ahmed, Shivam Singhal, Varsha Sharma, Sakyajit Bhattacharya, Aniruddha Sinha, Avik Ghose

IDENTIFICATION OF UTERINE CONTRACTIONS BY AN ENSEMBLE OF GAUSSIAN PROCESSES ........................................................................................................... 1020
   Liu Yang, Cassandra Heiselman, J. Gerald Quirk, Petar M. Djuric

ARRHYTHMIA CLASSIFICATION WITH HEARTBEAT-AWARE TRANSFORMER ................................................................. 1025
   Bin Wang, Chang Liu, Chuanyan Hu, Xudong Liu, Jun Cao

**BIO-3: MACHINE LEARNING FOR COVID-19 DIAGNOSIS**

MULTI-LEVEL GROUP TESTING WITH APPLICATION TO ONE-SHOT POOLED COVID-19 TESTS ................................................................. 1030
   Alejandro Cohen, Nir Shlezinger, Amit Solomon, Yonina C. Eldar, Muriel Médard

DETECTION OF COVID-19 THROUGH THE ANALYSIS OF VOCAL FOLD OSCILLATIONS ................................................................. 1035
   Mahmoud Al Ismail, Soham Deshmukh, Rita Singh

CT-CAPS: FEATURE EXTRACTION-BASED AUTOMATED FRAMEWORK FOR COVID-19 DISEASE IDENTIFICATION FROM CHEST CT SCANS USING CAPSULE NETWORKS ................................................................. 1040
   Shahin Heidarian, Parnian Afshar, Arash Moammadi, Moezedin Javad Rafiee Md, Anastasia Oikonomou Md, Konstantinos N. Plataniotis, Farnoosh Naderkhani

FEW-SHOT LEARNING FOR CT SCAN BASED COVID-19 DIAGNOSIS ........................................................................................................... 1045
   Yifan Jiang, Han Chen, Hanseok Ko, David K. Han
GRAPH-BASED PYRAMID GLOBAL CONTEXT REASONING WITH A SALIENCY-AWARE PROJECTION FOR COVID-19 LUNG INFECTIONS SEGMENTATION .......................................................... 1050
Huiimin Huang, Ming Cai, Lanfen Lin, Jing Zheng, Xiongwei Mao, Xiaohan Qian, Zhiyi Peng, Jianying Zhou, Yutaro Iwamoto, Xian-Hua Han, Yen-Wei Chen, Ruofeng Tong

INTERPRETING GLOTTAL FLOW DYNAMICS FOR DETECTING COVID-19 FROM VOICE ........ 1055
Soham Deshmukh, Mahmoud Al Ismail, Rita Singh

CYCLE GENERATIVE ADVERSARIAL NETWORK APPROACHES TO PRODUCE NOVEL PORTABLE CHEST X-RAYS IMAGES FOR COVID-19 DIAGNOSIS ................................................. 1060
Daniel I. Morís, Joaquim De Moura, Jorge Novo, Marcos Ortega

BIO-4: MACHINE LEARNING AND SIGNAL PROCESSING FOR NEURAL SIGNALS

EEG-BASED EMOTION CLASSIFICATION USING GRAPH SIGNAL PROCESSING .................. 1065
Seyed Saman Saboksayr, Gonzalo Mateos, Mujdat Cetin

GRANGER CAUSALITY BASED DIRECTIONAL PHASE-AMPLITUDE COUPLING MEASURE ............................................................. 1070
Tamanna T. K. Munia, Selin Aviyente

REPAC: RELIABLE ESTIMATION OF PHASE-AMPLITUDE COUPLING IN BRAIN NETWORKS ........................................................................................................ 1075
Giulia Cisotto

SUBSPACE ODDITY - OPTIMIZATION ON PRODUCT OF STIEFEL MANIFOLDS FOR EEG DATA ............................................................................................................. 1080
Maria Sayu Yamamoto, Florian Yger, Sylvain Chevallier

DECENTRALIZED MOTION INFERENCE AND REGISTRATION OF NEUROPIXEL DATA ...... 1085
Erdem Varol, Julien Boussard, Nishchal Dethe, Olivier Winter, Anne Urai, The International Brain Laboratory, Anne Churchland, Nick Steinmetz, Liam Paninski

DYNAMIC GRAPH LEARNING BASED ON GRAPH LAPLACIAN ........................................... 1090
Bo Jiang, Yiyi Yu, Hamid Krim, Spencer L. Smith

BIO-5: NEUROIMAGING AND NEURAL SIGNAL PROCESSING

MUTUAL INFORMATION FLOWS IN A BIVARIATE POINT PROCESS .................................. 1095
Syed Ahmed Pasha, Victor Solo

UNCERTAINTY-BASED BIOLOGICAL AGE ESTIMATION OF BRAIN MRI SCANS ............ 1100
Karim Armanious, Sherif Abdulatif, Wenbin Shi, Tobias Hepp, Sergios Gatidis, Bin Yang

SPARSE REPRESENTATION OF COMPLEX-VALUED FMRI DATA BASED ON HARD THRESHOLDING OF SPATIAL SOURCE PHASE ........................................................................ 1105
Jia-Yang Song, Miao-Ying Qi, Dun-Pei Lv, Chao-Ying Zhang, Qiu-Hua Lin, Vince D. Calhoun

TUCKER DECOMPOSITION FOR EXTRACTING SHARED AND INDIVIDUAL SPATIAL MAPS FROM MULTI-SUBJECT RESTING-STATE FMRI DATA ........................................ 1110
Yue Han, Qiu-Hua Lin, Li-Dan Kuang, Xiao-Feng Gong, Fengyu Cong, Vince D. Calhoun
RIEMANNIAN GEOMETRY-BASED DECODING OF THE DIRECTIONAL FOCUS OF AUDITORY ATTENTION USING EEG
Simon Geirnaert, Tom Francart, Alexander Bertrand

BIO-6: MEDICAL IMAGE SEGMENTATION
DFDM: A DEEP FEATURE DECOUPLING MODULE FOR LUNG NODULE SEGMENTATION
Wei Chen, Qiuli Wang, Sheng Huang, Xiaohong Zhang, Yucong Li, Chen Liu

PYRAMID U-_NET FOR RETINAL VESSEL SEGMENTATION
Jiawei Zhang, Yanchun Zhang, Xiaowei Xu

A PROBABILISTIC MODEL FOR SEGMENTATION OF AMBIGUOUS 3D LUNG NODULE
Xiaojiang Long, Wei Chen, Qiuli Wang, Xiaohong Zhang, Chen Liu, Yucong Li, Jiuquan Zhang

SEMI-SUPERVISED SKIN LESION SEGMENTATION WITH LEARNING MODEL CONFIDENCE
Zhiqiang Xie, Enmei Tu, Hao Zheng, Yun Gu, Jie Yang

A HYBRID FEATURE ENHANCEMENT METHOD FOR GL AND SEGMENTATION IN HISTOPATHOLOGY IMAGES
Xiangjiang Wu, Xiaoyi Li, Kai Hu, Zhineng Chen, Xieping Gao

AUTOMATED MULTI-ORGAN SEGMENTATION IN PET IMAGES USING CASCADED TRAINING OF A 3D U-NET AND CONVOLUTIONAL AUTOENCODER
Annika Liebgott, Charlotte Lorenz, Sergios Gatidis, Viet Chau Vu, Konstantin Nikolaou, Bin Yang

BIO-7: MEDICAL IMAGE FORMATION AND RECONSTRUCTION
IMPROVED SUPERVISED TRAINING OF PHYSICS-GUIDED DEEP LEARNING IMAGE RECONSTRUCTION WITH MULTI-MASKING
Burhaneddin Yaman, Seyed Amir Hossein Hosseini, Steen Moeller, Mehmet Akçakaya

FINE-GRAINED MRI RECONSTRUCTION USING ATTENTIVE SELECTION GENERATIVE ADVERSARIAL NETWORKS
Jingshuai Liu, Mehrdad Yaghoobi

ENSURE: ENSEMBLE STEIN’S UNBIASED RISK ESTIMATOR FOR UNSUPERVISED LEARNING
Hemant Kumar Aggarwal, Aniket Pramanik, Mathews Jacob

ULTRASOUND ELASTICITY IMAGING USING PHYSICS-BASED MODELS AND LEARNING-BASED PLUG-AND-PLAY PRIORS
Narges Mohammadi, Marvin M. Doyley, Mujdat Cetin

A PERIODIC FRAME LEARNING APPROACH FOR ACCURATE LANDMARK LOCALIZATION IN M-MODE ECHOCARDIOGRAPHY
Yinbing Tian, Shibiao Xu, Li Guo, Fuze Cong

A BIAS-REDUCING LOSS FUNCTION FOR CT IMAGE DENOISING
Madhuri Nagare, Roman Melnyk, Obaidullah Rahman, Ken D. Sauer, Charles A. Bouman
BIO-8: BIOLOGICAL IMAGE ANALYSIS

LEARNING BINARY SEMANTIC EMBEDDING FOR BREAST HISTOLOGY IMAGE CLASSIFICATION AND RETRIEVAL
Xiao Kang, Xingbo Liu, Xiushan Nie, Yilong Yin

CHANNEL ATTENTION RESIDUAL U-NET FOR RETINAL VESSEL SEGMENTATION
Changlu Guo, Márton Szemenyei, Yangtao Hu, Wenle Wang, Wei Zhou, Yugen Yi

CMIM: CROSS-MODAL INFORMATION MAXIMIZATION FOR MEDICAL IMAGING
Tristan Sylvain, Francis Dutil, Tess Berthier, Lisa Di Jorio, Margaux Luck, Devon Hjelm, Yoshua Bengio

STRUCTURE-ENHANCED ATTENTIVE LEARNING FOR SPINE SEGMENTATION FROM ULTRASOUND VOLUME PROJECTION IMAGES
Rui Zhao, Zixun Huang, Tianhuan Liu, Frank H. F. Leung, Sai Ho Ling, De Yang, Timothy Ting-Yan Lee, Daniel P. K. Lun, Yong-Ping Zheng, Kin-Man Lam

FOVEAL AVASCULAR ZONE SEGMENTATION OF OCTA IMAGES USING DEEP LEARNING APPROACH WITH UNSUPERVISED VESSEL SEGMENTATION
Zhijin Liang, Junkang Zhang, Cheolhong An

ACUTE LYMPHOBLASTIC LEUKEMIA DETECTION BASED ON ADAPTIVE UNSHARPENING AND DEEP LEARNING
Angelo Genovese, Mahdi S. Hosseini, Vincenzo Piuri, Konstantinos N. Plataniotis, Fabio Scotti

BIO-9: MEDICAL IMAGE ANALYSIS

META ORDINAL WEIGHTING NET FOR IMPROVING LUNG NODULE CLASSIFICATION
Yiming Lei, Hongming Shan, Junping Zhang

DEEPNODULE: MULTI-TASK LEARNING OF SEGMENTATION BOOTSTRAP FOR PULMONARY NODULE DETECTION
Jingqin Li, Kun Wang, Dan Yang, Xiaohong Zhang, Chen Liu

DENSE ATTENTION MODULE FOR ACCURATE PULMONARY NODULE DETECTION
Jiannan Liu, Jie Li, Fanyong Xue, Chentao Wu

UNSUPERVISED MULTIMODAL IMAGE REGISTRATION WITH ADAPTATIVE GRADIENT GUIDANCE
Zhe Xu, Jiangpeng Yan, Jie Luo, Xiufan Li, Jayender Jagadeesan

IMPROVING INTRAOPERATIVE LIVER REGISTRATION IN IMAGE-GUIDED SURGERY WITH LEARNING-BASED RECONSTRUCTION
Meng Jia, Matthew Kyan

A NEW FRAMEWORK BASED ON TRANSFER LEARNING FOR CROSS-DATABASE PNEUMONIA DETECTION
Xinxin Shan, Ying Wen
BIO-10: DEEP LEARNING FOR EEG ANALYSIS

HIERARCHICAL ATTENTION-BASED TEMPORAL CONVOLUTIONAL NETWORKS FOR EEG-BASED EMOTION RECOGNITION ................................................................. 1240
Chao Li, Boyang Chen, Ziping Zhao, Nicholas Cummins, Björn W. Schuller

DEEP MULTIWAY CANONICAL CORRELATION ANALYSIS FOR MULTI-SUBJECT EEG NORMALIZATION............................................................................ 1245
Jaswanth Reddy Katthi, Sriram Ganapathy

DYNAMIC GRAPH MODELING OF SIMULTANEOUS EEG AND EYE-TRACKING DATA FOR READING TASK IDENTIFICATION .............................................. 1250
Puneet Mathur, Trisha Mittal, Dinesh Manocha

LEARNING FROM HETEROGENEOUS EEG SIGNALS WITH DIFFERENTIABLE CHANNEL REORDERING .................................................................................... 1255
Aaqib Saeed, David Grangier, Olivier Pietquin, Neil Zeghidour

ENHANCING MULTI-CHANNEL EEG CLASSIFICATION WITH GRAMIAN TEMPORAL GENERATIVE ADVERSARIAL NETWORKS............................................ 1260
Chi Nok Enoch Kan, Richard J Povinelli, Dong Hye Ye

A NOVEL CONVOLUTIONAL NEURAL NETWORK MODEL TO REMOVE MUSCLE Artifacts FROM EEG ...................................................................................... 1265
Haoming Zhang, Chen Wei, Mingqi Zhao, Quanying Liu, Haiyan Wu

BIO-11: DEEP LEARNING FOR PHYSIOLOGICAL SIGNALS

MULTILABEL 12-LEAD ELECTROCARDIOGRAM CLASSIFICATION USING BEAT TO SEQUENCE AUTOENCODERS........................................................................... 1270
Alexander William Wong, Amir Salimi, Abram Hindle, Sunil Vasu Kalmady, Padma Kaul

CONTRASTIVE EMBEDDING LEARNING METHOD FOR RESPIRATORY SOUND CLASSIFICATION ............................................................................................ 1275
Wenjie Song, Jiqing Han, Hongwei Song

DECODING NEURAL REPRESENTATIONS OF RHYTHMIC SOUNDS FROM MAGNETOEENCEPHALOGRAPHY ................................................................. 1280
Pei-Chun Chang, Jia-Ren Chang, Po-Yu Chen, Li-Kai Cheng, Jen-Chuen Hsieh, Hsin-Yen Yu, Li-Fen Chen, Yong-Sheng Chen

LOW-DIMENSIONAL DENOISING EMBEDDING TRANSFORMER FOR ECG CLASSIFICATION........................................................................................................ 1285
Jian Guan, Wenbo Wang, Pengming Feng, Xinxin Wang, Wenwu Wang

SELF-SUPERVISED LEARNING FOR SLEEP STAGE CLASSIFICATION WITH PREDICTIVE AND DISCRIMINATIVE CONTRASTIVE CODING .......................... 1290
Qinfeng Xiao, Jing Wang, Jianan Ye, Hongjun Zhang, Yuyan Bu, Yiqiong Zhang, Hao Wu

LENGTH NO LONGER MATTERS: A REAL LENGTH ADAPTIVE ARRHYTHMIA CLASSIFICATION MODEL WITH MULTI-SCALE CONVOLUTION ........................................................................... 1295
Chuanqi Han, Fang Yu, Peng Wang, Ruoran Huang, Xi Huang, Li Cui
FEW-SHOT LEARNING FOR DECODING SURFACE ELECTROMYOGRAPHY FOR HAND GESTURE RECOGNITION
Elahe Rahimian, Soheil Zabihi, Amir Asif, S. Farokh Atashzar, Arash Mohammadi

BIO-12: FEATURE EXTRACTION AND FUSION FOR BIOMEDICAL APPLICATIONS

DEEPLUNG AUSCULTATION USING ACOUSTIC BIOMARKERS FOR ABNORMAL RESPIRATORY SOUND EVENT DETECTION
Upasana Tiwari, Swapnil Bhosale, Rupayan Chakraborty, Sunil Kumar Kopparapu

SPEAKER-INDEPENDENT BRAIN ENHANCED SPEECH DENOISING
Maryam Hosseini, Luca Celotti, Éric Plourde

SHAPELET BASED VISUAL ASSESSMENT OF CLUSTER TENDENCY IN ANALYZING COMPLEX UPPER LIMB MOTION
Shreyasi Datta, Chandan Karmakar, Punit Rathore, Marimuthu Palaniswami

HUMAN-CENTERED FAVORITE MUSIC CLASSIFICATION USING EEG-BASED INDIVIDUAL MUSIC PREFERENCE VIA DEEP TIME-SERIES CCA
Ryosuke Sawata, Takahiro Ogawa, Miki Haseyama

MULTI-SCALE AND MULTI-REGION FACIAL DISCRIMINATIVE REPRESENTATION FOR AUTOMATIC DEPRESSION LEVEL PREDICTION
Mingyue Niu, Jianhua Tao, Bin Liu

ECG HEART-BEAT CLASSIFICATION USING MULTIMODAL IMAGE FUSION
Zeeshan Ahmad, Anika Tabassum, Ling Guan, Naimul Khan

BIO-13: DEEP LEARNING FOR BIOMEDICAL APPLICATIONS

ESTIMATION OF VISUAL FEATURES OF VIEWED IMAGE FROM INDIVIDUAL AND SHARED BRAIN INFORMATION BASED ON FMRI DATA USING PROBABILISTIC GENERATIVE MODEL
Takaaki Higashi, Keisuke Maeda, Takahiro Ogawa, Miki Haseyama

HIERARCHICAL POSE CLASSIFICATION FOR INFANT ACTION ANALYSIS AND MENTAL DEVELOPMENT ASSESSMENT
Jianxiong Zhou, Zhongyu Jiang, Jang-Hee Yoo, Jenq-Neng Hwang

ON THE RELATIONSHIP BETWEEN SPEECH-BASED BREATHING SIGNAL PREDICTION EVALUATION MEASURES AND BREATHING PARAMETERS ESTIMATION
Zohreh Mostaani, Venkata Srikanth Nallanthighal, Aki Härmä, Helmer Strik, Mathew Magimai-Doss

PREDICTION OF EGFR MUTATION STATUS IN LUNG ADENOCARCINOMA USING MULTI-SOURCE FEATURE REPRESENTATIONS
Jianhong Cheng, Jin Liu, Meilin Jiang, Hailin Yue, Lin Wu, Jianxin Wang

TRAINING NEURAL NETWORKS WITH DOMAIN PATTERN-AWARE AUXILIARY TASK FOR SLEEP STAGING
Taeeheon Lee, Jeonghwan Hwang, Honggu Lee

CLASSIFICATION OF EXPERT-NOVICE LEVEL USING EYE TRACKING AND MOTION DATA VIA CONDITIONAL MULTIMODAL VARIATIONAL AUTOENCODER
Yusuke Akamatsu, Keisuke Maeda, Takahiro Ogawa, Miki Haseyama
CI-1: THEORY FOR COMPUTATIONAL IMAGING

GATE TRIMMING: ONE-SHOT CHANNEL PRUNING FOR EFFICIENT CONVOLUTIONAL NEURAL NETWORKS ......................................................... 1365
  Fang Yu, Chuanqi Han, Pengcheng Wang, Xi Huang, Li Cui

DEEP S^PR: SIMULTANEOUS SOURCE SEPARATION AND PHASE RETRIEVAL USING DEEP GENERATIVE MODELS ................................... 1370
  Christopher A. Metzler, Gordon Wetzstein

ADVERSARIAL ATTACKS ON OBJECT DETECTORS WITH LIMITED PERTURBATIONS ........ 1375
  Zhenbo Shi, Wei Yang, Zhenbo Xu, Zhi Chen, Yingjie Li, Haoran Zhu, Liusheng Huang

A CONSENSUS EQUILIBRIUM SOLUTION FOR DEEP IMAGE PRIOR POWERED BY RED ........ 1380
  Rakib Hyder, Hassan Mansour, Yanting Ma, Petros T. Boufounos, Pu Wang

SUREMAP: PREDICTING UNCERTAINTY IN CNN-BASED IMAGE RECONSTRUCTIONS USING STEIN'S UNBIASED RISK ESTIMATE ............ 1385
  Ruangrawee Kitichotkul, Christopher A. Metzler, Frank Ong, Gordon Wetzstein

MULTI-INITIALIZATION META-LEARNING WITH DOMAIN ADAPTATION ......................... 1390
  Zhengyu Chen, Donglin Wang

CI-2: COMPUTATIONAL IMAGING FOR INVERSE PROBLEMS

STOCHASTIC DEEP UNFOLDING FOR IMAGING INVERSE PROBLEMS ................................ 1395
  Jiaming Liu, Yu Sun, Weijie Gan, Xiaojian Xu, Brendt Wohlberg, Ulugbek S. Kamilov

FUSION-BASED DIGITAL IMAGE CORRELATION FRAMEWORK FOR STRAIN MEASUREMENT ................................................................. 1400
  Laixi Shi, Dehong Liu, Masaki Umeda, Norihiko Hana

LEARNING SPARSIFYING TRANSFORMS FOR IMAGE RECONSTRUCTION IN ELECTRICAL IMPEDANCE TOMOGRAPHY .............................. 1405
  Kaiyi Yang, Narong Borijindargoon, Boon Poh Ng, Saiprasad Ravishankar, Bihan Wen

D-VDAMP: DENOISING-BASED APPROXIMATE MESSAGE PASSING FOR COMPRESSIVE MRI ........................................................................... 1410
  Christopher A. Metzler, Gordon Wetzstein

EMPIRICALLY ACCELERATING SCALED GRADIENT PROJECTION USING DEEP NEURAL NETWORK FOR INVERSE PROBLEMS IN IMAGE PROCESSING ...................................... 1415
  Byung Hyun Lee, Se Young Chun

SYNTHETIC APERTURE ACOUSTIC IMAGING WITH DEEP GENERATIVE MODEL BASED SOURCE DISTRIBUTION PRIOR ........................................... 1420
  Boqiang Fan, Samarjit Das

CI-3: COMPUTATIONAL PHOTOGRAPHY

NON-LOCAL SINGLE IMAGE DE-RAINING WITHOUT DECOMPOSITION ................................ 1425
  Chaobing Zheng, Zhengguo Li, Yuwen Li, Shiqian Wu
FRAME-RATE-AWARE AGGREGATION FOR EFFICIENT VIDEO SUPER-RESOLUTION .......... 1430
Takashi Isobe, Fang Zhu, Shengjin Wang

MEASUREMENT CODING FRAMEWORK WITH ADJACENT PIXELS BASED
MEASUREMENT MATRIX FOR COMPRESSIVELY SENSED IMAGES .............................................. 1435
Rentao Wan, Jinjia Zhou, Bowen Huang, Hui Zeng, Yibo Fan

MULTIVIEW SENSING WITH UNKNOWN PERMUTATIONS: AN OPTIMAL TRANSPORT
APPROACH...................................................................................................................................................  1440
Yanting Ma, Petros T. Boufounos, Hassan Mansour, Shuchin Aeron

A HIGH-FRAME-RATE EYE-TRACKING FRAMEWORK FOR MOBILE DEVICES ......................... 1445
Yuhu Chang, Changyang He, Yingying Zhao, Tun Lu, Ning Gu

CATILOC: CAMERA IMAGE TRANSFORMER FOR INDOOR LOCALIZATION.............................. 1450
Ali Ghofrani, Rahil Mahdian Toroghi, Seyed Mojtaba Tabatabaei

CI-4: REMOTE SENSING AND CODED APERTURE IMAGING

SAR IMAGE AUTOFOCUSED USING WIRTINGER CALCULUS AND CAUCHY
REGULARIZATION ......................................................................................................................................  1455
Zi-Yao Zhang, Odysseas Pappas, Alin Achim

A HOMOGENEITY-BASED MULTISCALE HYPERSPECTRAL IMAGE REPRESENTATION
FOR SPARSE SPECTRAL UNMIXING ...................................................................................................... 1460
Luciano C. Ayres, Sérgio J. M. De Almeida, José C. M. Bermudez, Ricardo A. Borsoi

LEARNING TO ESTIMATE KERNEL SCALE AND ORIENTATION OF DEFOCUS BLUR
WITH ASYMMETRIC CODED APERTURE .............................................................................................. 1465
Jisheng Li, Qi Dai, Jiangtao Wen

TRANSMITTANCE REGULARIZER FOR BINARY CODED APERTURE DESIGN IN A
COMPUTATIONAL IMAGING END-TO-END APPROACH .................................................................... 1470
Jorge Bacca, Tatiana Gelvez, Henry Arguello

FOURIER TRANSFORMATION AUTOENCODERS FOR ANOMALY DETECTION ......................... 1475
Demetris Lappas, Vasileios Argyriou, Dimitrios Makris

ZERO-GRADIENT CONSTRAINTS FOR DESTRIPIING OF REMOTE-SENSING DATA .................. 1480
Kazuki Naganuma, Saori Takeyama, Shunsuke Ono

IVMSP-1: OBJECT DETECTION 1

MULTI-SCALE SAMPLE SELECTION BASED ON STATISTICAL CHARACTERISTICS FOR
OBJECT DETECTION .................................................................................................................................. 1485
Zhiguo Li, Yuan Yuan, Dandan Ma

MS-CSPN: MULTI-SCALE CASCADE SPATIAL PYRAMID NETWORK FOR OBJECT
DETECTION.................................................................................................................................................. 1490
Tianyuan Wang, Can Ma, Haoshan Su, Weiping Wang

DUAL-STREAM NETWORK BASED ON GLOBAL GUIDANCE FOR SALIENT OBJECT
DETECTION.................................................................................................................................................. 1495
Shuyong Gao, Qianyu Guo, Wei Zhang, Wenyang Zhang, Zhongwei Ji
SSFENET: SPATIAL AND SEMANTIC FEATURE ENHANCEMENT NETWORK FOR OBJECT DETECTION .................................................................................................................................. 1500
Tianyuan Wang, Can Ma, Haoshan Su, Weiping Wang

SALIENCY-DRIVEN VERSATILE VIDEO CODING FOR NEURAL OBJECT DETECTION .......... 1505
Kristian Fischer, Felix Fleckenstein, Christian Herglotz, André Kaup

OBJECT-ORIENTED RELATIONAL DISTILLATION FOR OBJECT DETECTION ...................... 1510
Shuyu Miao, Rui Feng

IVMSP-2: OBJECT DETECTION 2

ENSEMBLING OBJECT DETECTORS FOR IMAGE AND VIDEO DATA ANALYSIS................. 1515
Kateryna Chumachenko, Jenni Raitoharju, Alexandros Iosifidis, Moncef Gabbouj

TRAINING REAL-TIME PANORAMIC OBJECT DETECTORS WITH VIRTUAL DATASET ......... 1520
Qing-Yang Shen, Tian-Guo Huang, Peng-Xin Ding, Jia He

FAST: FEATURE AGGREGATION FOR DETECTING SALIENT OBJECT IN REAL-TIME .......... 1525
Lv Tang, Bo Li, Yanliang Wu, Bo Xiao, Shouhong Ding

EXPLOITING THE DUAL-TREE COMPLEX WAVELET TRANSFORM FOR SHIP WAKE DETECTION IN SAR IMAGERY ................................................................................................................ 1530
Wanli Ma, Alin Achim, Oktay Karakus

TASK-RELATED SELF-SUPERVISED LEARNING FOR REMOTE SENSING IMAGE CHANGE DETECTION ................................................................................................................................. 1535
Zhinan Cai, Zhiyu Jiang, Yuan Yuan

UNSUPERVISED COMMON PARTICULAR OBJECT DISCOVERY AND LOCALIZATION BY ANALYZING A MATCH GRAPH ........................................................................................................ 1540
Makoto Okuda, Shin'Ichi Satoh, Yoichi Sato, Yutaka Kidawara

IVMSP-3: IMAGE & VIDEO CODING 1

PREDICTIVE CODING FOR LOSSLESS DATASET COMPRESSION........................................... 1545
Madeleine Barowsky, Alexander Mariona, Flavio P. Calmon

ADAPTIVE DUAL TREE STRUCTURE FOR SCREEN CONTENT CODING ................................. 1550
Weijia Zhu, Jizheng Xu, Li Zhang, Yue Wang

SNR-ADAPTIVE DEEP JOINT SOURCE-CHANNEL CODING FOR WIRELESS IMAGE TRANSMISSION .............................................................................................................................................. 1555
Mingze Ding, Jiahui Li, Mengyao Ma, Xiaopeng Fan

RELYING ON A RATE CONSTRAINT TO REDUCE MOTION ESTIMATION COMPLEXITY........ 1560
Gabriel B. Sant’Anna, Luiz Henrique Cancellier, Ismael Seidel, Mateus Grellert, José Luis Güntzel

A NOVEL VIEWPORT-ADAPTIVE MOTION COMPENSATION TECHNIQUE FOR FISHEYE VIDEO ................................................................................................................................. 1565
Andy Regensky, Christian Herglotz, André Kaup
RATE-DISTORTION OPTIMIZED MOTION ESTIMATION FOR ON-THE-SPHERE COMPRESSION OF 360 VIDEOS ................................................................. 1570
Alban Marie, Navid Mahmoudian Bidgoli, Thomas Maugéy, Aline Roumy

IVMSP-4: IMAGE & VIDEO CODING 2

ADAPTIVE GOP SIZE DECISION FOR MULTI-PASS VIDEO CODING BASED ON HIDDEN MARKOV MODEL .............................................................. 1575
Bohan Li, Jingning Han, Yaowu Xu

IMPROVED INTRA MODE CODING BEYOND AV1 ......................................................... 1580
Yize Jin, Liang Zhao, Xin Zhao, Shan Liu, Alan C. Bovik

DECISION TREE BASED INTER PARTITION TERMINATION FOR AV1 ENCODING .............. 1585
Xinyao Chen, Yiwei Zhang, Yanghao Li, Jiangtao Wen

IMAGE CODING FOR MACHINES: AN END-TO-END LEARNED APPROACH .................. 1590
Nam Le, Honglei Zhang, Francesco Cricri, Ramin Ghaznavi-Youvalari, Esa Rahtu

SPARSE FLOW ADVERSARIAL MODEL FOR ROBUST IMAGE COMPRESSION ................. 1595
Shihui Zhao, Shuyuan Yang, Zhi Liu, Zhixi Feng, Xu Liu

HVS-BASED PERCEPTUAL COLOR COMPRESSION OF IMAGE DATA .......................... 1600
Lee Prangnell, Victor Sanchez

IVMSP-5: SUPER-RESOLUTION 1

HOCA: HIGHER-ORDER CHANNEL ATTENTION FOR SINGLE IMAGE SUPER-RESOLUTION ......................................................................................... 1605
Yalei Lv, Tao Dai, Bin Chen, Jian Lu, Shu-Tao Xia, Jingchao Cao

IMAGE SUPER-RESOLUTION USING MULTI-RESOLUTION ATTENTION NETWORK ............ 1610
Anqi Liu, Sumei Li, Yongli Chang

REAL IMAGE SUPER-RESOLUTION USING TOKEN BASED CONTEXTUAL ATTENTION ....... 1615
Zhihong Pan, Baopu Li

FEATURE REDUNDANCY MINING: DEEP LIGHT-WEIGHT IMAGE SUPER-RESOLUTION MODEL ...................................................................................... 1620
Jun Xiao, Wenqi Jia, Kin-Man Lam

LIGHTWEIGHT NON-LOCAL NETWORK FOR IMAGE SUPER-RESOLUTION .................... 1625
Risheng Wang, Tao Lei, Wenzheng Zhou, Qi Wang, Hongying Meng, Asoke K. Nandi

LIGHTWEIGHT AND ACCURATE SINGLE IMAGE SUPER-RESOLUTION WITH CHANNEL SEGREGATION NETWORK .............................................. 1630
Zhong-Han Niu, Xi-Peng Lin, An-Ni Yu, Yang-Hao Zhou, Yu-Bin Yang

IVMSP-6: SUPER-RESOLUTION 2 & MULTI-SCALE PROCESSING

DEEP LEARNING ARCHITECTURAL DESIGNS FOR SUPER-RESOLUTION OF NOISY IMAGES ......................................................................................... 1635
Angel Villar-Corrales, Franziska Schirrmacher, Christian Riess
JOINT COUPLED TRANSFORM LEARNING FRAMEWORK FOR MULTIMODAL IMAGE SUPER-RESOLUTION .......................................................... 1640
   Andrew Gigie, A Anil Kumar, Angshul Majumdar, Kriti Kumar, M Girish Chandra

HYPERSPECTRAL IMAGE SUPER-RESOLUTION VIA ADJACENT SPECTRAL FUSION STRATEGY ....................................................................................... 1645
   Qiang Li, Qi Wang, Xuelong Li

RAW DATA PROCESSING FOR PRACTICAL TIME-OF-FLIGHT SUPER-RESOLUTION ................. 1650
   Miguel Heredia Conde

EDGE-AWARE MULTI-SCALE PROGRESSIVE COLORIZATION ........................................................................................................... 1655
   Jun Xia, Guanghua Tan, Yi Xiao, Fangqiang Xu, Chi-Sing Leung

LEARNING REPRESENTATION OF MULTI-SCALE OBJECT FOR FINE-GRAINED IMAGE RETRIEVAL .................................................................................. 1660
   Kangbo Sun, Jie Zhu

IVMSP-7: MACHINE LEARNING FOR IMAGE PROCESSING

SUPER-RESOLUTION AND INFECTION EDGE DETECTION CO-GUIDED LEARNING FOR COVID-19 CT SEGMENTATION .................................................................................................................... 1665
   Yu Sang, Jinguang Sun, Simiao Wang, Heng Qi, Keqiu Li

GATING FEATURE DENSE NETWORK FOR SINGLE ANISOTROPIC MR IMAGE SUPER-RESOLUTION .................................................................................................................. 1670
   Weidong He, Yangjinan Hu, Lulu Wang, Zhongshi He, Jinglong Du

ADAPTABLE ENSEMBLE DISTILLATION.................................................................................................................. 1675
   Yankai Wang, Dawei Yang, Wei Zhang, Zhe Jiang, Wenqiang Zhang

A SCALE INVARIANT MEASURE OF FLATNESS FOR DEEP NETWORK MINIMA ............ 1680
   Akshay Rangamani, Nam H. Nguyen, Abhishek Kumar, Dzung Phan, Sang Peter Chin, Trac D. Tran

MULTI-ORDER ADVERSARIAL REPRESENTATION LEARNING FOR COMPOSED QUERY IMAGE RETRIEVAL .......................................................... 1685
   Zhixiao Fu, Xinyuan Chen, Jianfeng Dong, Shouling Ji

DEEP NEURAL NETWORKS WITH FLEXIBLE COMPLEXITY WHILE TRAINING BASED ON NEURAL ORDINARY DIFFERENTIAL EQUATIONS ...................................................................... 1690
   Zhengbo Luo, Sei-Ichiro Kamata, Zitang Sun, Weilian Zhou

IVMSP-8: MACHINE LEARNING FOR IMAGE PROCESSING

IMPROVING MEMORY BANKS FOR UNSUPERVISED LEARNING WITH LARGE MINI-BATCH, CONSISTENCY AND HARD NEGATIVE MINING ........................................................................ 1695
   Adrian Bulat, Enrique Sánchez-Lozano, Georgios Tzimiropoulos

ROBUST BINARY LOSS FOR MULTI-CATEGORY CLASSIFICATION WITH LABEL NOISE ............................................................................................................. 1700
   Defu Liu, Guowu Yang, Jinzhao Wu, Jiayi Zhao, Fengmao Lv
A PLUG AND PLAY FAST INTERSECTION OVER UNION LOSS FOR BOUNDARY BOX REGRESSION .......................................................... 1705
  Zengsheng Kuang, Xian Fang, Ruixun Zhang, Xiuli Shao, Hongpeng Wang

ATTRIBUTE DECOMPOSITION FOR FLOW-BASED DOMAIN MAPPING ............................................. 1710
  Sheng-Jhe Huang, Jen-Tzung Chien

ADA-SISE: ADAPTIVE SEMANTIC INPUT SAMPLING FOR EFFICIENT EXPLANATION OF CONVOLUTIONAL NEURAL NETWORKS .................................................. 1715
  Mahesh Sudhakar, Sam Sattarzadeh, Konstantinos N. Plataniotis, Jongseong Jang, Yeonjeong Jeong, Hyunwoo Kim

NETWORK PRUNING USING LINEAR DEPENDENCY ANALYSIS ON FEATURE MAPS ............ 1720
  Hao Pan, Zhongdi Chao, Jiang Qian, Bojin Zhuang, Shaojun Wang, Jing Xiao

IVMSP-9: ZERO AND FEW SHORT LEARNING

MULTIPLE-INPUT MULTIPLE-OUTPUT FUSION NETWORK FOR GENERALIZED ZERO-SHOT LEARNING .......................................................... 1725
  Fangming Zhong, Guangze Wang, Zhikui Chen, Xu Yuan, Feng Xia

REPRESENTATIVE LOCAL FEATURE MINING FOR FEW-SHOT LEARNING ............................................. 1730
  Kun Yan, Lingbo Liu, Jun Hou, Ping Wang

KAN: KNOWLEDGE-AUGMENTED NETWORKS FOR FEW-SHOT LEARNING ............................................. 1735
  Zeyang Zhu, Xin Lin

FEW-SHOT IMAGE CLASSIFICATION WITH MULTI-FACET PROTOTYPES .................................................. 1740
  Kun Yan, Zied Bouraoui, Ping Wang, Shoaib Jameel, Steven Schockaert

SELF-SUPERVISED LEARNING FOR FEW-SHOT IMAGE CLASSIFICATION ............................................. 1745
  Da Chen, Yuefeng Chen, Yuhong Li, Feng Mao, Yuan He, Hui Xue

DOMAIN ADAPTATION FOR LEARNING GENERATOR FROM PAIRED FEW-SHOT DATA ........ 1750
  Chun-Chih Teng, Pin-Yu Chen, Wei-Chen Chiu

IVMSP-10: METRIC LEARNING AND INTERPRETABILITY

DEEP SEMI-SUPERVISED METRIC LEARNING VIA IDENTIFICATION OF MANIFOLD MEMBERSHIPS .......................................................... 1755
  Furen Zhuang, Pierre Moulin

A RANKED SIMILARITY LOSS FUNCTION WITH PAIR WEIGHTING FOR DEEP METRIC LEARNING ............................................................................. 1760
  Jian Wang, Zhichao Zhang, Dongmei Huang, Wei Song, Quanmiao Wei, Xinyue Li

STATISTICAL DISTANCE METRIC LEARNING FOR IMAGE SET RETRIEVAL .................................................. 1765
  Ting-Yao Hu, Alexander G. Hauptmann

DISTRIBUTION-AWARE HIERARCHICAL WEIGHTING METHOD FOR DEEP METRIC LEARNING ............................................................................. 1770
  Yinong Zhu, Yong Feng, Mingliang Zhou, Baohua Qiang, Leong Hou U, Jiajie Zhu
INTEGRATED GRAD-CAM: SENSITIVITY-AWARE VISUAL EXPLANATION OF DEEP
CONVOLUTIONAL NETWORKS VIA INTEGRATED GRADIENT-BASED SCORING ...................... 1775
Sam Sattarzadeh, Mahesh Sudhakar, Konstantinos N. Plataniotis, Jongseong Jang, Yeonjeong
Jeong, Hyunwoo Kim

VISUALIZING ASSOCIATION IN EXEMPLAR-BASED CLASSIFICATION ............................... 1780
Taiga Kashima, Ryuichiro Hataya, Hideki Nakayama

IVMSP-11: IMAGE & VIDEO SEGMENTATION

HFGCNET: HIGH-FREQUENCY GRAPH REASONING FOR FINER SEMANTIC IMAGE
SEGMENTATION .......................................................................................................................... 1785
Zitang Sun, Ruojing Wang, Zhengbo Luo, Weili Chen

UNSUPERVISED IMAGE SEGMENTATION WITH SPATIAL TRIplet MARKOV TREES .......... 1790
Hugo Gangloff, Jean-Baptiste Courbot, Emmanuel Monfrini, Christophe Collet

CROSS SCENE VIDEO FOREGROUND SEGMENTATION VIA CO-OCCURRENCE
PROBABILITY ORIENTED SUPERVISED AND UNSUPERVISED MODEL INTERACTION ...... 1795
Dong Liang, Bin Kang, Xinyu Liu, Han Sun, Liyan Zhang, Ningzhong Liu

INSTANCE SEGMENTATION WITH THE NUMBER OF CLUSTERS INCORPORATED IN
EMBEDDING LEARNING .............................................................................................................. 1800
Jianfeng Cao, Hong Yan

DECouple THE HIGH-FREQUENCY AND LOW-FREQUENCY INFORMATION OF
IMAGES FOR SEMANTIC SEGMENTATION .............................................................................. 1805
Lianlei Shan, Xiaobin Li, Weiqiang Wang

MPDNET: A 3D MISSING PART DETECTION NETWORK BASED ON POINT CLOUD
SEGMENTATION ......................................................................................................................... 1810
Zhaoxin Fan, Hongyan Liu, Jun He, Min Zhang, Xiaoyong Du

IVMSP-12: IMAGE & VIDEO INTERPRETATION AND UNDERSTANDING

SM+: REFINED SCALE MATCH FOR TINY PERSON DETECTION ........................................... 1815
Nan Jiang, Xuehui Yu, Xiaoke Peng, Yuqi Gong, Zhenjun Han

SUB-BAND GROUPING SPECTRAL FEATURE-ATTENTION BLOCK FOR
HYPERSPECTRAL IMAGE CLASSIFICATION .............................................................................. 1820
Weilian Zhou, Sei-Ichiro Kamata, Zhengbo Luo

UNSUPERVISED STACKED CAPSULE AUTOENCODER FOR HYPERSPECTRAL IMAGE
CLASSIFICATION ......................................................................................................................... 1825
Erting Pan, Yong Ma, Xiaoguang Mei, Fan Fan, Jiayi Ma

ROBUST GRAPH AUTOENCODER FOR HYPERSPECTRAL ANOMALY DETECTION .......... 1830
Ganghui Fan, Yong Ma, Jun Huang, Xiaoguang Mei, Jiayi Ma

IVMSP-13: IMAGE ENHANCEMENT AND RESTORATION

REFLECTANCE-ORIENTED PROBABILISTIC EQUALIZATION FOR IMAGE
ENHANCEMENT .......................................................................................................................... 1835
Xiaomeng Wu, Yongqiang Sun, Akisato Kimura, Kunio Kashino
PD-GAN: PERCEPTUAL-DETAILS GAN FOR EXTREMELY NOISY LOW LIGHT IMAGE ENHANCEMENT ................................................................. 1840
Yijun Liu, Zhengning Wang, Yi Zeng, Hao Zeng, Deming Zhao

HETEROGENEOUS TWO-STREAM NETWORK WITH HIERARCHICAL FEATURE PREFUSION FOR MULTISPECTRAL PAN-SHARPENING ................................................................. 1845
Dong Wang, Yunpeng Bai, Bendi Bai, Chanyue Wu, Ying Li

SYNERGIC FEATURE ATTENTION FOR IMAGE RESTORATION ........................................................................................................... 1850
Chong Mou, Jian Zhang

EFFICIENT MULTI-OBJECTIVE GANS FOR IMAGE RESTORATION ........................................................................................................ 1855
Jingwen Su, Hujun Yin

SELF-CONVOLUTION: A HIGHLY-EFFICIENT OPERATOR FOR NON-LOCAL IMAGE RESTORATION ........................................................................................................................ 1860
Lanqing Guo, Zhiyuan Zha, Saiprasad Ravishankar, Bihan Wen

IVMSP-14: HYPERSPECTRAL IMAGING

NMF-SAE: AN INTERPRETABLE SPARSE AUTOENCODER FOR HYPERSPECTRAL UNMIXING ................................................................................................................................. 1865
Fengchao Xiong, Jun Zhou, Minchao Ye, Jianfeng Lu, Yuntao Qian

AN ADMM BASED NETWORK FOR HYPERSPECTRAL UNMIXING TASKS ........................................................................................................... 1870
Chao Zhou, Miguel R. D. Rodrigues

VARIATIONAL AUTOENCODERS FOR HYPERSPECTRAL UNMIXING WITH ENDMEMBER VARIABILITY ........................................................................................................... 1875
Shuaikai Shi, Min Zhao, Lijun Zhang, Jie Chen

AUGMENTED GAUSSIAN LINEAR MIXTURE MODEL FOR SPECTRAL VARIABILITY IN HYPERSPECTRAL UNMIXING ........................................................................................................... 1880
Yaser Esmaeili Salehani, Ehsan Arabnejad, Saeed Gazor

UTDN: AN UNSUPERVISED TWO-STREAM DIRICHLET-NET FOR HYPERSPECTRAL UNMIXING ................................................................................................................................. 1885
Qiwen Jin, Yong Ma, Xiaoguang Mei, Hao Li, Jiayi Ma

LAPLACIAN REGULARIZED TENSOR LOW-RANK MINIMIZATION FOR HYPERSPECTRAL SNAPSHOT COMPRESSIVE IMAGING ........................................................................................................... 1890
Yi Yang, Fei Jiang, Hongtao Lu

IVMSP-15: LOCAL DESCRIPTORS AND TEXTURE

COMPRESSIONING LOCAL DESCRIPTOR MODELS FOR MOBILE APPLICATIONS ........................................................................................................... 1895
Roy Miles, Krystian Mikolajczyk

VK-NET: CATEGORY-LEVEL POINT CLOUD REGISTRATION WITH UNSUPERVISED ROTATION INVARIANT KEYPOINTS ........................................................................................................... 1900
Zhi Chen, Wei Yang, Zhenbo Xu, Zhenbo Shi, Liusheng Huang

MATCHING AS COLOR IMAGES: THERMAL IMAGE LOCAL FEATURE DETECTION AND DESCRIPTION ........................................................................................................... 1905
Bhavesh Deshpande, Sourabh Hanamsheth, Yawen Lu, Guoyu Lu
FRAME RATE UP-CONVERSION USING KEY POINT AGNOSTIC FREQUENCY-SELECTIVE MESH-TO-GRID RESAMPLING ........................................................... 1910
   Viktoria Heimann, Andreas Spruck, André Kaup

EFFICIENT REAL-TIME VIDEO STABILIZATION WITH A NOVEL LEAST SQUARES FORMULATION ................................................................. 1915
   Jianwei Ke, Alex J Watras, Jae-Jun Kim, Hewei Liu, Hongrui Jiang, Yu Hen Hu

DECOMPOSING TEXTURES USING EXPONENTIAL ANALYSIS ........................................................... 1920
   Yuan Hou, Annie Cuyt, Wen-Shin Lee, Deepayan Bhowmik

IVMSP-16: POINT CLOUDS AND DEPTH

G-ARRAYS: GEOMETRIC ARRAYS FOR EFFICIENT POINT CLOUD PROCESSING ................. 1925
   Hoda Roodaki, Masoud Dehyadegari, Mahdi Nazm Bojnordi

QOE-DRIVEN AND TILE-BASED ADAPTIVE STREAMING FOR POINT CLOUDS ................. 1930
   Lisha Wang, Chenglin Li, Wenrui Dai, Junni Zou, Hongkai Xiong

DYNAMIC POINT CLOUD COMPRESSION USING A CUBOID ORIENTED DISCRETE COSINE BASED MOTION MODEL ........................................ 1935
   Ashik Ahmed, Manoranjan Paul, Manzur Murshed, David Taubman

AN ADAPTIVE PYRAMID SINGLE-VIEW DEPTH LOOKUP TABLE CODING METHOD ............ 1940
   Yangang Cai, Ronggang Wang, Song Gu, Jian Zhang, Wen Gao

PATCH DECODER-SIDE DEPTH ESTIMATION IN MPEG IMMERSIVE VIDEO ...................... 1945
   Marta Milovanovic, Félix Henry, Marco Cagnazzo, Joël Jung

GEOMETRY CONSISTENCY OF AUGMENTED REALITY BASED ON SEMANTICS .................. 1950
   Hongyan Quan, Mingwei Yao, Xiaoxiao Qian

IVMSP-17: LOOKING AT PEOPLE

WHAT AND WHERE TO FOCUS IN PERSON SEARCH ............................................................ 1955
   Tong Zhou, Kun Tian

STABLE AND EFFECTIVE ONE-STEP METHOD FOR PERSON SEARCH ............................... 1960
   Ning Lv, Xuezhi Xiang, Xinyao Wang, Jie Yang, Rokia Abdeen, Abdulmotaleb El Saddik

AN ADAPTIVE PART-BASED MODEL FOR PERSON RE-IDENTIFICATION .............................. 1965
   Xi-Peng Lin, Yu-Bin Yang

CROWD COUNTING VIA MULTI-LEVEL REGRESSION WITH LATENT GAUSSIAN MAPS .......... 1970
   Yukang Gao, Hua Yang

LIGHTWEIGHT DUAL-TASK NETWORKS FOR CROWD COUNTING IN AERIAL IMAGES ....... 1975
   Ye Tian, Chengzhen Duan, Ruijin Zhang, Zhiwei Wei, Hongpeng Wang

SANET++: ENHANCED SCALE AGGREGATION WITH DENSELY CONNECTED FEATURE FUSION FOR CROWD COUNTING ........................................... 1980
   Siyang Pan, Yanyun Zhao, Fei Su, Zhicheng Zhao
IVMSP-18: FACES IN IMAGES & VIDEOS

ATTENTIVE SEMANTIC EXPLORING FOR MANIPULATED FACE DETECTION ........................................... 1985
Zehao Chen, Hua Yang

EFFICIENT FACE MANIPULATION VIA DEEP FEATURE DISENTANGLEMENT AND REINTEGRATION NET ............................................................................................................................... 1990
Bin Cheng, Tao Dai, Bin Chen, Shutao Xia, Xiu Li

CONTINUOUS FACE AGING GENERATIVE ADVERSARIAL NETWORKS ....................................................... 1995
Seogkyu Jeon, Pilhyeon Lee, Kibeom Hong, Hyeran Byun

FAST INVERSE MAPPING OF FACE GANS ..................................................................................................... 2000
Nicky Bayat, Vahid Reza Khazaie, Yalda Mohsenzadeh

MULTI-LEVEL ADAPTIVE REGION OF INTEREST AND GRAPH LEARNING FOR FACIAL ACTION UNIT RECOGNITION .................................................................................................................. 2005
Jingwei Yan, Boyuan Jiang, Jingjing Wang, Qiang Li, Chunmao Wang, Shiliang Pu

BRIDGING UNPAIRED FACIAL PHOTOS AND SKETCHES BY LINE-DRAWINGS .................................... 2010
Meimei Shang, Fei Gao, Xiang Li, Jingjie Zhu, Lingna Dai

IVMSP-19: DERAINEING AND DEHAZING

TEMPORAL RAIN DECOMPOSITION WITH SPATIAL STRUCTURE GUIDANCE FOR VIDEO DERAINEING ........................................................................................................................................ 2015
Xinwei Xue, Ying Ding, Long Ma, Yi Wang, Risheng Liu, Xin Fan

GTA-NET: GRADUAL TEMPORAL AGGREGATION NETWORK FOR FAST VIDEO DERAINEING .......................................................................................................................................................... 2020
Xinwei Xue, Xiangyu Meng, Long Ma, Risheng Liu, Xin Fan

DENSE FEATURE PYRAMID GRIDS NETWORK FOR SINGLE IMAGE DERAINEING .................................... 2025
Zhen Wang, Cong Wang, Zhixun Su, Junyang Chen

A FAST AND EFFICIENT NETWORK FOR SINGLE IMAGE DERAINEING .................................................. 2030
Youzhao Yang, Hong Lu

DNANET: DENSE NESTED ATTENTION NETWORK FOR SINGLE IMAGE DEHAZING .................................. 2035
Dongdong Ren, Jinbao Li, Meng Han, Minglei Shu

FWB-NET: FRONT WHITE BALANCE NETWORK FOR COLOR SHIFT CORRECTION IN SINGLE IMAGE DEHAZING VIA ATMOSPHERIC LIGHT ESTIMATION .................................................. 2040
Cong Wang, Yan Huang, Yuexian Zou, Yong Xu

IVMSP-20: DENOISING AND DEBLURRING

LEARNING INTEGRODIFFERENTIAL MODELS FOR IMAGE DENOISING .................................................. 2045
Tobias Alt, Joachim Weickert

UNROLLING OF DEEP GRAPH TOTAL VARIATION FOR IMAGE DENOISING ........................................... 2050
Huy Vu, Gene Cheung, Yonina C. Eldar
LEARNING MODEL-BLIND TEMPORAL DENOISERS WITHOUT GROUND TRUTHS
Yanghao Li, Bichuan Guo, Jiangtao Wen, Zhen Xia, Shan Liu, Yuxing Han

IMAGE DENOISING BASED ON CORRELATION ADAPTIVE SPARSE MODELING
Hangfan Liu, Jian Zhang, Chong Mou

NASA: A NOISE-ADAPTIVE AND STRUCTURE-AWARE LEARNING FRAMEWORK FOR IMAGE DEBLURRING
Xiaokun Liu, Long Ma, Risheng Liu, Wei Zhong, Xin Fan, Zhongxuan Luo

MULTIPLE AUXILIARY NETWORKS FOR SINGLE BLIND IMAGE DEBLURRING
Chen Li, Qi Wang, Shaoteng Liu, Xuelong Li

IVMSP-21: IMAGE & VIDEO QUALITY

JOINT LEARNING OF IMAGE AESTHETIC QUALITY ASSESSMENT AND SEMANTIC RECOGNITION BASED ON FEATURE ENHANCEMENT
Xiangfei Liu, Xiushan Nie, Zhen Shen, Yilong Yin

NESTED ERROR MAP GENERATION NETWORK FOR NO-REFERENCE IMAGE QUALITY ASSESSMENT
Junming Chen, Haiqiang Wang, Ge Li, Shan Liu

REGRESSION OR CLASSIFICATION? NEW METHODS TO EVALUATE NO-REFERENCE PICTURE AND VIDEO QUALITY MODELS
Zhengzhong Tu, Chia-Ju Chen, Li-Heng Chen, Yilin Wang, Neil Birkbeck, Balu Adsumilli, Alan C. Bovik

BLIND IMAGE QUALITY EVALUATOR WITH SCALE ROBUSTNESS
Ci Wang, Mei Li

MULTI-SCALE FEATURE-GUIDED STEREOSCOPIC VIDEO QUALITY ASSESSMENT BASED ON 3D CONVOLUTIONAL NEURAL NETWORK
Yingjie Feng, Sumei Li, Yongli Chang

NO-REFERENCE STEREOSCOPIC IMAGE QUALITY ASSESSMENT BASED ON THE HUMAN VISUAL SYSTEM
Fan Meng, Sumei Li, Yongli Chang

IVMSP-22: IMAGE & VIDEO SENSING, MODELING AND REPRESENTATION

STEREO RECTIFICATION BASED ON EPIPOLAR CONSTRAINED NEURAL NETWORK
Yuxing Wang, Yawen Lu, Guoyu Lu

MULTI-SCALE CASCADE DISPARITY REFINEMENT STEREOSTEREO NETWORK
Xiaogang Jia, Wei Chen, Zhengfa Liang, Xin Luo, Mingfei Wu, Yusong Tan, Libo Huang

HIERARCHICAL CONTEXT GUIDED AGGREGATION NETWORK FOR STEREO MATCHING
Jun Peng, Wangduo Xie, Zijing Huang, Wei Chen, Yong Zhao

COST AFFINITY LEARNING NETWORK FOR STEREO MATCHING
Shenglun Chen, Baopu Li, Wei Wang, Hong Zhang, Haojie Li, Zhihui Wang
VIDEO QUALITY PREDICTION USING VOXEL-WISE FMRI MODELS OF THE VISUAL CORTEX ................................................................. 2125
   Naga Sailaja Mahankali, Sumohana S. Channappayya

TENSOR DECOMPOSITION VIA CORE TENSOR NETWORKS ................................................................................................. 2130
   Jianfu Zhang, Zerui Tao, Liqing Zhang, Qibin Zhao

IVMSP-23: APPLICATIONS 1

SIGN LANGUAGE SEGMENTATION WITH TEMPORAL CONVOLUTIONAL NETWORKS .......... 2135
   Katrin Renz, Nicolaj C. Stache, Samuel Albanie, Gül Varol

AN ADAPTIVE DISCRIMINANT AND SPARSITY FEATURE DESCRIPTOR FOR FINGER VEIN RECOGNITION ................................................................. 2140
   Shuyi Li, Bob Zhang

ROUTINGGAN: ROUTING AGE PROGRESSION AND REGRESSION WITH DISENTANGLLED LEARNING ........................................................................ 2145
   Zhizhong Huang, Junping Zhang, Hongming Shan

SEMANTIC-AWARE UNPAIRED IMAGE-TO-IMAGE TRANSLATION FOR URBAN SCENE IMAGES ........................................................................ 2150
   Zongyao Li, Ren Togo, Takahiro Ogawa, Miki Haseyama

FONTNET: ON-DEVICE FONT UNDERSTANDING AND PREDICTION PIPELINE .................. 2155
   Rakshith S, Rishabh Khurana, Vibhav Agarwal, Jayesh Rajkumar Vachhani, Guggilla Bhanodai

AGENT-ENVIRONMENT NETWORK FOR TEMPORAL ACTION PROPOSAL GENERATION .................................................................................. 2160
   Viet-Khoa Vo-Ho, Ngan Le, Kashu Kamazaki, Akihiro Sugimoto, Minh-Triet Tran

IVMSP-24: APPLICATIONS 2

ADAPTIVE MULTI-DOMAIN LEARNING FOR OUTDOOR 3D HUMAN POSE AND SHAPE ESTIMATION ................................................................. 2165
   Zhaoyang Gui, Shanshan Zhang, Kangkan Wang, Jian Yang, Pong Chi Yuen

LIGHTWEIGHT HUMAN POSE ESTIMATION UNDER RESOURCE-LIMITED SCENES .......... 2170
   Zhe Zhang, Jie Tang, Gangshan Wu

ABSOLUTE 3D POSE ESTIMATION AND LENGTH MEASUREMENT OF SEVERELY DEFORMED FISH FROM MONOCULAR VIDEOS IN LONGLINE FISHING ........................................................................ 2175
   Jie Mei, Jenq-Neng Hwang, Suzanne Romain, Craig Rose, Braden Moore, Kelsey Magrane

CAMERA CALIBRATION WITH POSE GUIDANCE .................................................................................. 2180
   Yuzhuo Ren, Feng Hu

REAL VERSUS FAKE 4K - AUTHENTIC RESOLUTION ASSESSMENT .................................. 2185
   Rishi Rajesh Shah, Vyas Anirudh Akundy, Zhou Wang

PERCEPTUAL QUALITY ASSESSMENT FOR RECOGNIZING TRUE AND PSEUDO 4K CONTENT .................................................................................. 2190
   Wenhan Zhu, Guangtao Zhai, Xiongkuo Min, Xiaokang Yang, Xiao-Ping Zhang
IVMSP-25: TRACKING

A NEW TUBULAR STRUCTURE TRACKING ALGORITHM BASED ON CURVATURE-PENALIZED PERCEPTUAL GROUPING .................................................................................................. 2195
   Li Liu, Da Chen, Minglei Shu, Huazhong Shu, Laurent D. Cohen

MULTIPLE HUMAN TRACKING IN NON-SPECIFIC COVERAGE WITH WEARABLE CAMERAS............................................................................................................................................. 2200
   Sibo Wang, Ruize Han, Wei Feng, Song Wang

FINE-GRAINED POSE TEMPORAL MEMORY MODULE FOR VIDEO POSE ESTIMATION AND TRACKING...................................................................................................................................... 2205
   Chaoyi Wang, Yang Hua, Tao Song, Zhengui Xue, Ruhui Ma, Neil Robertson, Haibing Guan

DRAWING ORDER RECOVERY FROM TRAJECTORY COMPONENTS ............................................ 2210
   Minghao Yang, Xukang Zhou, Yangchang Sun, Jinglong Chen, Baohua Qiang

DEEP HASHING FOR MOTION CAPTURE DATA RETRIEVAL .......................................................... 2215
   Na Lv, Ying Wang, Zhiquan Feng, Jingliang Peng

HIERARCHICAL ATTENTION FUSION FOR GEO-LOCALIZATION ................................................. 2220
   Liqi Yan, Yiming Cui, Yingjie Chen, Dongfang Liu

IVMSP-26: ATTENTION FOR VISION

ATTENTIONLITE: TOWARDS EFFICIENT SELF-ATTENTION MODELS FOR VISION..................... 2225
   Souvik Kundu, Sairam Sundaresan

ATTENTION-GUIDED SECOND-ORDER POOLING CONVOLUTIONAL NETWORKS.......................... 2230
   Shannan Chen, Qiule Sun, Cunhua Li, Jianxin Zhang, Qiang Zhang

SA-NET: SHUFFLE ATTENTION FOR DEEP CONVOLUTIONAL NEURAL NETWORKS.................. 2235
   Qing-Long Zhang, Yu-Bin Yang

AN ATTENTION BASED WAVELET CONVOLUTIONAL MODEL FOR VISUAL SALIENCY DETECTION........................................................................................................................................... 2240
   Reshmi S. Bhooshan, K Suresh

CASCADE ATTENTION FUSION FOR FINE-GRAINED IMAGE CAPTIONING BASED ON
MULTI-LAYER LSTM ................................................................................................................................. 2245
   Shuang Wang, Yun Meng, Yu Gu, Lei Zhang, Xiutiao Ye, Jingxian Tian, Licheng Jiao

WEBLY SUPERVISED DEEP ATTENTIVE QUANTIZATION ................................................................. 2250
   Jinpeng Wang, Bin Chen, Tao Dai, Shu-Tao Xia

IVMSP-27: MULTI-MODAL SIGNAL PROCESSING

UNSUPERVISED AUDIO-VISUAL SUBSPACE ALIGNMENT FOR HIGH-STAKES DECEPTION DETECTION.................................................................................................................. 2255
   Leena Mathur, Maja J. Mataric

VIOLENCE DETECTION IN VIDEOS BASED ON FUSING VISUAL AND AUDIO INFORMATION .......................................................... 2260
   Wen-Feng Pang, Qian-Hua He, Yong-Jian Hu, Yan-Xiong Li
QUERYD: A VIDEO DATASET WITH HIGH-QUALITY TEXT AND AUDIO NARRATIONS .......... 2265
Andreea-Maria Oncescu, João F. Henriques, Yang Liu, Andrew Zisserman, Samuel Albanie

GENERATING NATURAL QUESTIONS FROM IMAGES FOR MULTIMODAL ASSISTANTS .......... 2270
Alkesh Patel, Akanksha Bindal, Hadas Kotek, Christopher Klein, Jason Williams

AN ADAPTIVE MULTI-SCALE AND MULTI-LEVEL FEATURES FUSION NETWORK
WITH PERCEPTUAL LOSS FOR CHANGE DETECTION ......................................................... 2275
Jialang Xu, Yang Luo, Xinyue Chen, Chunbo Luo

SEEHEAR: SIGNER DIARISATION AND A NEW DATASET ........................................... 2280
Samuel Albanie, Giil Varol, Liliane Momeni, Triantafyllos Afouras, Andrew Brown, Chuhan
Zhang, Ernesto Coto, Necati Cihan Camgöz, Ben Saunders, Abhishek Dutta, Neil Fox,
Richard Bowden, Bencie Woll, Andrew Zisserman

IVMSP-28: IMAGE SYNTHESIS

SEMANTIC IMAGE SYNTHESIS FROM INACCURATE AND COARSE MASKS .................. 2285
Kai Katsumata, Hideki Nakayama

RANGE GUIDED DEPTH REFINEMENT AND UNCERTAINTY-AWARE AGGREGATION
FOR VIEW SYNTHESIS ........................................................................................................ 2290
Yuan Chang, Yisong Chen, Guoping Wang

DP-VTON: TOWARD DETAIL-PRESERVING IMAGE-BASED VIRTUAL TRY-ON
NETWORK .............................................................................................................................. 2295
Yuan Chang, Tao Peng, Ruhan He, Xinxong Hu, Junping Liu, Zili Zhang, Minghua Jiang

LIGHT FIELD STYLE TRANSFER WITH LOCAL ANGULAR CONSISTENCY .................. 2300
Dónal Egan, Martin Alain, Aljosa Smolic

SKIP ATTENTION GAN FOR REMOTE SENSING IMAGE SYNTHESIS ......................... 2305
Kai Deng, Kun Zhang, Ping Yao, Siyuan Cheng, Peng He

IMAGE GENERATION BASED ON TEXTURE GUIDED VAE-AGAN FOR REGIONS OF
INTEREST DETECTION IN REMOTE SENSING IMAGES ................................................ 2310
Libao Zhang, Yanan Liu

IVMSP-29: SEMANTIC SEGMENTATION

EADNET: EFFICIENT ASYMMETRIC DILATED NETWORK FOR SEMANTIC
SEGMENTATION .................................................................................................................. 2315
Qihang Yang, Tao Chen, Jiayuan Fan, Ye Lu, Chongyan Zuo, Qinghua Chi

LTAF-NET: LEARNING TASK-AWARE ADAPTIVE FEATURES AND REFINING MASK
FOR FEW-SHOT SEMANTIC SEGMENTATION .................................................................... 2320
Binjie Mao, Lingfeng Wang, Shiming Xiang, Chunhong Pan

CGAN-NET: CLASS-GUIDED ASYMMETRIC NON-LOCAL NETWORK FOR REAL-TIME
SEMANTIC SEGMENTATION .............................................................................................. 2325
Hanlin Chen, Qingyong Hu, Jangang Yang, Jing Wu, Yulan Guo

AGGREGATION ARCHITECTURE AND ALL-TO-ONE NETWORK FOR REAL-TIME
SEMANTIC SEGMENTATION ............................................................................................. 2330
Kuntao Cao, Xi Huang, Jie Shao
<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>NLKD: USING COARSE ANNOTATIONS FOR SEMANTIC SEGMENTATION BASED ON</td>
<td>2335</td>
</tr>
<tr>
<td>KNOWLEDGE DISTILLATION</td>
<td></td>
</tr>
<tr>
<td>Dong Liang, Yun Du, Han Sun, Liyan Zhang, Ningzhong Liu, Mingqiang</td>
<td></td>
</tr>
<tr>
<td>Wei</td>
<td></td>
</tr>
<tr>
<td>KNOWLEDGE REASONING FOR SEMANTIC SEGMENTATION</td>
<td>2340</td>
</tr>
<tr>
<td>Shengjia Chen, Zhixin Li, Xiwei Yang</td>
<td></td>
</tr>
<tr>
<td><strong>IVMSP-30: INVERSE PROBLEMS IN IMAGE &amp; VIDEO PROCESSING</strong></td>
<td></td>
</tr>
<tr>
<td>NON-CONVEX SPARSE DEVIATION MODELING VIA GENERATIVE MODELS</td>
<td>2345</td>
</tr>
<tr>
<td>Yaxi Yang, Hailin Wang, Haiquan Qiu, Jianjun Wang, Yao Wang</td>
<td></td>
</tr>
<tr>
<td>IMRNET: AN ITERATIVE MOTION COMPENSATION AND RESIDUAL</td>
<td>2350</td>
</tr>
<tr>
<td>RECONSTRUCTION NETWORK FOR VIDEO COMPRESSED SENSING</td>
<td></td>
</tr>
<tr>
<td>Xin Yang, Chunling Yang</td>
<td></td>
</tr>
<tr>
<td>DEEP COLOR CONSTANCY USING TEMPORAL GRADIENT UNDER AC LIGHT SOURCES</td>
<td>2355</td>
</tr>
<tr>
<td>Jeong-Won Ha, Jun-Sang Yoo, Jong-Ok Kim</td>
<td></td>
</tr>
<tr>
<td>END-TO-END LEARNING OF VARIATIONAL MODELS AND SOLVERS FOR THE</td>
<td>2360</td>
</tr>
<tr>
<td>RESOLUTION OF INTERPOLATION PROBLEMS</td>
<td></td>
</tr>
<tr>
<td>R. Fablet, L. Drumetz, F. Rousseau</td>
<td></td>
</tr>
<tr>
<td>MULTI-MODELS FUSION FOR LIGHT FIELD ANGULAR SUPER-RESOLUTION</td>
<td>2365</td>
</tr>
<tr>
<td>Fengyin Cao, Ping An, Xinpeng Huang, Chao Yang, Qiang Wu</td>
<td></td>
</tr>
<tr>
<td>HIDE CHOPIN IN THE MUSIC: EFFICIENT INFORMATION STEGANOGRAPHY VIA</td>
<td>2370</td>
</tr>
<tr>
<td>RANDOM SHUFFLING</td>
<td></td>
</tr>
<tr>
<td>Zhun Sun, Chao Li, Qibin Zhao</td>
<td></td>
</tr>
<tr>
<td><strong>IVMSP-31: APPLICATIONS 3</strong></td>
<td></td>
</tr>
<tr>
<td>POINTER NETWORKS FOR ARBITRARY-SHAPED TEXT SPOTTING</td>
<td>2375</td>
</tr>
<tr>
<td>Yi Zhang, Wei Yang, Zhenbo Xu, Yingjie Li, Zhi Chen, Liusheng Huang</td>
<td></td>
</tr>
<tr>
<td>ROTATION INVARIANCE ANALYSIS OF LOCAL CONVOLUTIONAL FEATURES IN</td>
<td>2380</td>
</tr>
<tr>
<td>IMAGE RETRIEVAL</td>
<td></td>
</tr>
<tr>
<td>Longjiao Zhao, Yu Wang, Jien Kato</td>
<td></td>
</tr>
<tr>
<td>SIGNATURE FEATURE MARKING ENHANCED IRM FRAMEWORK FOR DRONE IMAGE</td>
<td>2385</td>
</tr>
<tr>
<td>ANALYSIS IN PRECISION AGRICULTURE</td>
<td></td>
</tr>
<tr>
<td>Atharva Kadethankar, Neelam Sinha, Vinayaka Hegde, Abhishek Burman</td>
<td></td>
</tr>
<tr>
<td>VEHICLE 3D LOCALIZATION IN ROAD SCENES VIA A MONOCULAR MOVING</td>
<td>2390</td>
</tr>
<tr>
<td>CAMERA</td>
<td></td>
</tr>
<tr>
<td>Yanting Zhang, Aotian Zheng, Ke Han, Yizhou Wang, Jenq-Neng Hwang</td>
<td></td>
</tr>
<tr>
<td>GPS-DENIED NAVIGATION USING SAR IMAGES AND NEURAL NETWORKS</td>
<td>2395</td>
</tr>
<tr>
<td>Teresa White, Jesse Wheeler, Colton Lindstrom, Randall Christensen,</td>
<td></td>
</tr>
<tr>
<td>Kevin R. Moon</td>
<td></td>
</tr>
<tr>
<td>ATTENTION-EMBEDDED DECOMPOSED NETWORK WITH UNPAIRED CT IMAGES</td>
<td>2400</td>
</tr>
<tr>
<td>PRIOR FOR METAL ARTIFACT REDUCTION</td>
<td></td>
</tr>
<tr>
<td>Binyu Zhao, Qianqian Ren, Jinbao Li, Yafeng Zhao</td>
<td></td>
</tr>
</tbody>
</table>
IVMSP-32: APPLICATIONS 4

PARTIAL FEATURE AGGREGATION NETWORK FOR REAL-TIME OBJECT COUNTING ........................................ 2405
Houshun Yu, Li Zhang

A BAYESIAN INFERENCE APPROACH FOR LOCATION-BASED MICRO MOTIONS USING RADIO FREQUENCY SENSING .................................................................................................................. 2410
David A. Maluf, Amr Elnakeeb, Matt Silverman

ROBUST SPATIAL-TEMPORAL CORRELATION MODEL FOR BACKGROUND INITIALIZATION IN SEVERE SCENE ........................................................................................................................................ 2415
Yuheng Deng, Wenjun Zhou, Bo Peng, Dong Liang, Shun’ichi Kaneko

2D-FRFT BASED FREQUENCY SHIFT-IN Variant DIGITAL IMAGE ENCRYPTION ........................................ 2420
Lei Gao, Lin Qi, Ling Guan

CAPTURING BANDING IN IMAGES: DATABASE CONSTRUCTION AND OBJECTIVE ASSESSMENT ......................................................................................................................................................... 2425
Akshay Kapoor, Jatin Sapra, Zhou Wang

ON THE CAMERA POSITION DITHERING IN VISUAL 3D RECONSTRUCTION .............................................. 2430
Qier An, Yuan Shen

IVMSP-33: ACTION RECOGNITION

LONG-SHORT TEMPORAL MODELING FOR EFFICIENT ACTION RECOGNITION ..................................... 2435
Liyu Wu, Yuexian Zou, Can Zhang

MULTI-DIRECTIONAL CONVOLUTION NETWORKS WITH SPATIAL-TEMPORAL FEATURE PYRAMID MODULE FOR ACTION RECOGNITION .............................................................................. 2440
Bohong Yang, Zijian Wang, Wu Ran, Hong Lu, Yi-Ping Phoebe Chen

UNSUPERVISED MOTION REPRESENTATION ENHANCED NETWORK FOR ACTION RECOGNITION ......................................................................................................................................................... 2445
Xiaohang Yang, Lingtong Kong, Jie Yang

AN IMPROVED DEEP RELATION NETWORK FOR ACTION RECOGNITION IN STILL IMAGES .................. 2450
Wei Wu, Jiale Yu

HUMAN-AWARE COARSE-TO-FINE ONLINE ACTION DETECTION .................................................................. 2455
Zichen Yang, Di Huang, Jie Qin, Yunhong Wang

SRF-NET: SELECTIVE RECEPTIVE FIELD NETWORK FOR ANCHOR-FREE TEMPORAL ACTION DETECTION .......................................................................................................................................... 2460
Ranyu Ning, Can Zhang, Yuexian Zou

IVMSP-34: INPAINTING AND OCCLUSIONS HANDLING

SEMANTIC-AWARE CONTEXT AGGREGATION FOR IMAGE INPAINTING .................................................. 2465
Zhilin Huang, Chujun Qin, Ruixin Liu, Zhenyu Weng, Yuesheng Zhu

BISHIFT-NET FOR IMAGE INPAINTING ........................................................................................................ 2470
Xue Zhou, Tao Dai, Yong Jiang, Shu-Tao Xia
OAS-NET: OCCLUSION AWARE SAMPLING NETWORK FOR ACCURATE OPTICAL FLOW

Lingtong Kong, Xiaohang Yang, Jie Yang

MASK4D: 4D CONVOLUTION NETWORK FOR LIGHT FIELD OCCLUSION REMOVAL

Yingjie Li, Wei Yang, Zhenbo Xu, Zhi Chen, Zhenbo Shi, Yi Zhang, Liusheng Huang

SELF-SUPERVISED DEPTH ESTIMATION VIA IMPLICIT CUES FROM VIDEOS

Jianrong Wang, Ge Zhang, Zhenyu Wu, Xuewei Li, Li Liu

SCENE COMPLETENESS-AWARE LIDAR DEPTH COMPLETION FOR DRIVING SCENARIO

Cho-Ying Wu, Ulrich Neumann

IFS-1: MULTIMEDIA FORENSICS 1

SEMI-SUPERVISED FEATURE EMBEDDING FOR DATA SANITIZATION IN REAL-WORLD EVENTS

Bahram Lavi, José Nascimento, Anderson Rocha

EXPOSING GAN-GENERATED FACES USING INCONSISTENT CORNEAL SPECULAR HIGHLIGHTS

Shu Hu, Yuezun Li, Siwei Lyu

A FEATURES DECOUPLING METHOD FOR MULTIPLE MANIPULATIONS IDENTIFICATION IN IMAGE OPERATION CHAINS

Jiaxin Chen, Xin Liao, Wei Wang, Zheng Qin

SUBJECTIVE AND OBJECTIVE EVALUATION OF DEEPFAKE VIDEOS

Pavel Korshunov, Sébastien Marcel

FORENSICABILITY OF DEEP NEURAL NETWORK INFERENCCE PIPELINES

Alexander Schlögl, Tobias Kupek, Rainer Böhme

IFS-2: MULTIMEDIA FORENSICS 2

SERN: STANCE EXTRACTION AND REASONING NETWORK FOR FAKE NEWS DETECTION

Jianhui Xie, Song Liu, Ruixin Liu, Yinghong Zhang, Yuesheng Zhu

AN EFFICIENT PAPER ANTI-COUNTERFEITING METHOD BASED ON MICROSTRUCTURE ORIENTATION ESTIMATION

Yuhao Sun, Xin Liao, Jianfeng Liu

LEARNING DOUBLE-COMPRESSION VIDEO FINGERPRINTS LEFT FROM SOCIAL-MEDIA PLATFORMS

Irene Amerini, Aris Anagnostopoulos, Luca Maiano, Lorenzo Ricciardi Celsi

CHECKING PRNU USABILITY ON MODERN DEVICES

C. Albisani, M. Iuliani, A. Piva

HANDWRITTEN DIGITS RECONSTRUCTION FROM UNLABELLED EMBEDDINGS

Thomas Thebaud, Gaël Le Lan, Anthony Larcher
EFFECT OF VIDEO PIXEL-BINNING ON SOURCE ATTRIBUTION OF MIXED MEDIA.................. 2545
    Samet Taspinar, Manoranj Mohanty, Nasir Memon

IFS-3: FORENSICS AND BIOMETRICS

COMBINING DYNAMIC IMAGE AND PREDICTION ENSEMBLE FOR CROSS-DOMAIN
FACE ANTI-SPOOFING ................................................................. 2550
    Lingling Lv, Youjun Xiang, Xianfeng Li, Hanye Huang, Rongju Ruan, Xiaoyan Xu, Yuli Fu
LABEL-GUIDED DICTIONARY PAIR LEARNING FOR ECG BIOMETRIC RECOGNITION ........ 2555
    Mingzhu Ma, Gongping Yang, Kuikui Wang, Yuwen Huang, Yilong Yin
BACKDOOR ATTACK AGAINST SPEAKER VERIFICATION............................................. 2560
    Tongqing Zhai, Yiming Li, Ziqi Zhang, Baoyuan Wu, Yong Jiang, Shu-Tao Xia
CLASS-CONDITIONAL DEFENSE GAN AGAINST END-TO-END SPEECH ATTACKS .......... 2565
    Mohammad Esmaeilpour, Patrick Cardinal, Alessandro Lameiras Koerich
SELFGAIT: A SPATIOTEMPORAL REPRESENTATION LEARNING METHOD FOR SELF-
SUPERVISED GAIT RECOGNITION .................................................. 2570
    Yiqun Liu, Yi Zeng, Jian Pu, Hongming Shan, Peiyang He, Junping Zhang
ATTACK ON PRACTICAL SPEAKER VERIFICATION SYSTEM USING UNIVERSAL
ADVERSARIAL PERTURBATIONS .......................................................... 2575
    Weiyi Zhang, Shuning Zhao, Le Liu, Jianmin Li, Xingliang Cheng, Thomas Fang Zheng,
    Xiaolin Hu

IFS-4: SURVEILLANCE, BIOMETRICS AND SECURITY

HIGHLY EFFICIENT PROTECTION OF BIOMETRIC FACE SAMPLES WITH SELECTIVE
JPEG2000 ENCRYPTION .............................................................. 2580
    Heinz Hofbauer, Yoanna Martínez-Díaz, Simon Kirchgasser, Heydi Méndez-Vázquez,
    Andreas Uhl
DEEP AUTO-ENCODING AND BIOHASHING FOR SECURE FINGER VEIN RECOGNITION.... 2585
    Hatef Otroshi Shahreza, Sébastien Marcel
TOPIC SEQUENCE EMBEDDING FOR USER IDENTITY LINKAGE FROM
HETEROGENEOUS BEHAVIOR DATA ................................................... 2590
    Jinzhu Yang, Wei Zhou, Wanhui Qian, Jizhong Han, Songlin Hu
LOOKING THROUGH WALLS: INFERRING SCENES FROM VIDEO-SURVEILLANCE
ENCRYPTED TRAFFIC ........................................................................ 2595
    Daniele Mari, Samuele Giuliano Piazzetta, Sara Bordin, Luca Pajola, Sebastiano Verde,
    Simone Milani, Mauro Conti
OPTIMAL ATTACKING STRATEGY AGAINST ONLINE REPUTATION SYSTEMS WITH
CONSIDERATION OF THE MESSAGE-BASED PERSUASION PHENOMENON .................. 2600
    Zhanjiang Chen, H. Vicky Zhao
IFS-5: PRIVACY AND INFORMATION SECURITY

STEP-GAN: A ONE-CLASS ANOMALY DETECTION MODEL WITH APPLICATIONS TO POWER SYSTEM SECURITY ................................................................. 2605
Mohammad Adiban, Arash Safari, Giampiero Salvi

APPLICATION-LAYER DDOS ATTACKS WITH MULTIPLE EMULATION DICTIONARIES ........ 2610
Michele Cirillo, Mario Di Mauro, Vincenzo Matta, Marco Tambasco

SECRET KEY GENERATION OVER WIRELESS CHANNELS USING SHORT BLOCKLENGTH MULTILEVEL SOURCE POLAR CODING ................................................................. 2615
Henri Hentilä, Yanina Y. Shkel, Visa Koivunen

EFFICIENT NETWORK PROTECTION GAMES AGAINST MULTIPLE TYPES OF STRATEGIC ATTACKERS .......................................................................................................................... 2620
Zhifan Xu, Melike Baykal-Gürsoy

DETECTION OF MALICIOUS DNS AND WEB SERVERS USING GRAPH-BASED APPROACHES ................................................................................................................................. 2625
Jinyuan Jia, Zheng Dong, Jie Li, Jack W. Stokes

LOW COMPLEXITY SECURE P-TENSOR PRODUCT COMPRESSED SENSING RECONSTRUCTION OUTSOURCING AND IDENTITY AUTHENTICATION IN CLOUD .................. 2630
Mengdi Wang, Di Xiao, Jia Liang

IFS-6: ANONYMIZATION, SECURITY AND PRIVACY

PRIVACY-PRESERVING NEAR NEIGHBOR SEARCH VIA SPARSE CODING WITH AMBIGUATION ................................................................................................................................. 2635
Behrooz Razeghi, Sohrab Ferdowski, Dimche Kostadinov, Flavio P. Calmon, Slava Voloshynovskiy

PRIVACY-PRESERVING OPTIMAL INSULIN DOSING DECISION ................................................................. 2640
Zuobin Ying, Shuanglong Cao, Shengmin Xu, Ximeng Liu, Lingjuan Lyu, Cen Chen, Li Wang

PRIVACY-ACCURACY TRADE-OFF OF INFERENCE AS SERVICE ................................................................................................................................. 2645
Yulu Jin, Lifeng Lai

FEDERATED LEARNING WITH LOCAL DIFFERENTIAL PRIVACY: TRADE-OFFS BETWEEN PRIVACY, UTILITY, AND COMMUNICATION ................................................................................................. 2650
Muah Kim, Onur Günlü, Rafael F. Schaefer

SCALABLE PRIVACY-PRESERVING DISTRIBUTED EXTREMELY RANDOMIZED TREES FOR STRUCTURED DATA WITH MULTIPLE COLLUDING PARTIES ................................................................................................................................. 2655
Amin Aminifar, Fazeel Rabbi, Yngve Lamo

ACTIVE PRIVACY-UTILITY TRADE-OFF AGAINST A HYPOTHESIS TESTING ADVERSARY ................................................................................................................................. 2660
Ecenaz Erdemir, Pier Luigi Dragotti, Deniz Gündüz
IFS-7: INFORMATION HIDING, CRYPTOGRAPHY AND CYBERSECURITY

BAITRADAR: A MULTI-MODEL CLICKBAIT DETECTION ALGORITHM USING DEEP LEARNING ............................................................................................................................ 2665
Bhanuka Gamage, Adnan Labib, Aisha Joomun, Chern Hong Lim, Koksheik Wong

ENABLING EFFICIENT AND EXPRESSIVE SPATIAL KEYWORD QUERIES ON ENCRYPTED DATA ................................................................. 2670
Xiangyu Wang, Jianfeng Ma, Ximeng Liu

PRIVACY-PRESERVING CLOUD-BASED DNN INFEERENCE ......................................................... 2675
Shangyu Xie, Bingyu Liu, Yuan Hong

CRYPTO-ORIENTED NEURAL ARCHITECTURE DESIGN ................................................................. 2680
Avital Shafran, Gil Segev, Shmuel Peleg, Yedid Hoshen

INTEGRATING DEEP LEARNING WITH FIRST-ORDER LOGIC PROGRAMMED CONSTRAINTS FOR ZERO-DAY PHISHING ATTACK DETECTION ........................................................................ 2685
Seok-Jun Bu, Sung-Bae Cho

IMPROVED PROBABILISTIC CONTEXT-FREE GRAMMARS FOR PASSWORDS USING WORD EXTRACTION .............................................................................................. 2690
Haibo Cheng, Wenting Li, Ping Wang, Kaitai Liang

IFS-8: WATERMARKING AND DATA HIDING

ENHANCING IMAGE STEGANOGRAPHY VIA STEGO GENERATION AND SELECTION .................. 2695
Tingting Song, Minglin Liu, Weiqi Luo, Peijia Zheng

SYNCHRONOUS MULTI-BIT AUDIO WATERMARKING BASED ON PHASE SHIFTING ................. 2700
Shengbei Wang, Weitao Yuan, Zhen Zhang, Jianming Wang, Masashi Unoki

IMAGE STEGANOGRAPHY BASED ON ITERATIVE ADVERSARIAL PERTURBATIONS ONTO A SYNCHRONIZED-DIRECTIONS SUB-IMAGE ......................................................... 2705
Xinghong Qin, Shunquan Tan, Weixuan Tang, Bin Li, Jiwu Huang

EXTENDING THE REVERSE JPEG COMPATIBILITY ATTACK TO DOUBLE COMPRESSED IMAGES ............................................................................................. 2710
Jan Butora, Jessica Fridrich

REVERSIBLE DATA HIDING IN JPEG IMAGES FOR PRIVACY PROTECTION ......................... 2715
Yuxuan Huang, Xin Cao, Hao-Tian Wu, Yiu-Ming Cheung

A LAYERED EMBEDDING-BASED SCHEME TO COPE WITH INTRA-FRAME DISTORTION DRIFT IN IPM-BASED HEVC STEGANOGRAPHY .......................................................... 2720
Xiaojing Jia, Jie Wang, Yongliang Liu, Xiangui Kang, Yun-Qing Shi

MLSP-1: DEEP LEARNING TRAINING METHODS

META-LEARNING WITH ATTENTION FOR IMPROVED FEW-SHOT LEARNING .......................... 2725
Zejiang Hou, Anwar Walid, Sun-Yuan Kung
<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>B-SMALL: A BAYESIAN NEURAL NETWORK APPROACH TO SPARSE MODEL-AGNOSTIC META-LEARNING</td>
<td>2730</td>
</tr>
<tr>
<td>Anish Madan, Ranjitha Prasad</td>
<td></td>
</tr>
<tr>
<td>DEEP TRANSFORM AND METRIC LEARNING NETWORKS</td>
<td>2735</td>
</tr>
<tr>
<td>Wen Tang, Emilie Chouzenoux, Jean-Christophe Pesquet, Hamid Krim</td>
<td></td>
</tr>
<tr>
<td>ROBUSTNESS AND DIVERSITY SEEKING DATA-FREE KNOWLEDGE DISTILLATION</td>
<td>2740</td>
</tr>
<tr>
<td>Pengchao Han, Jihong Park, Shiqiang Wang, Yejun Liu</td>
<td></td>
</tr>
<tr>
<td>ENSEMBLE DISTILLATION APPROACHES FOR GRAMMATICAL ERROR CORRECTION</td>
<td>2745</td>
</tr>
<tr>
<td>Y. Fathullah, M. J. F. Gales, A. Malinin</td>
<td></td>
</tr>
<tr>
<td>TRAIN YOUR CLASSIFIER FIRST: CASCADE NEURAL NETWORKS TRAINING FROM UPPER LAYERS TO LOWER LAYERS</td>
<td>2750</td>
</tr>
<tr>
<td>Shucong Zhang, Cong-Thanh Do, Rama Doddipatla, Erfan Loweimi, Peter Bell, Steve Renals</td>
<td></td>
</tr>
<tr>
<td><strong>MLSP-2: DEEP LEARNING TRAINING METHODS 2</strong></td>
<td></td>
</tr>
<tr>
<td>HOW CONVOLUTIONAL NEURAL NETWORKS DEAL WITH ALIASING</td>
<td>2755</td>
</tr>
<tr>
<td>Antônio H. Ribeiro, Thomas B. Schön</td>
<td></td>
</tr>
<tr>
<td>CANET: CONTEXT-AWARE LOSS FOR DESCRIPTOR LEARNING</td>
<td>2760</td>
</tr>
<tr>
<td>Tianyou Chen, Xiaoguang Hu, Jin Xiao, Guofeng Zhang, Hui Ruan</td>
<td></td>
</tr>
<tr>
<td>PROGRESSIVE MULTI-STAGE FEATURE MIX FOR PERSON RE-IDENTIFICATION</td>
<td>2765</td>
</tr>
<tr>
<td>Yan Zhang, Binyu He, Li Sun, Qingli Li</td>
<td></td>
</tr>
<tr>
<td>USING DEEP IMAGE PRIORS TO GENERATE COUNTERFACTUAL EXPLANATIONS</td>
<td>2770</td>
</tr>
<tr>
<td>Vivek Narayanaswamy, Jayaraman J. Thiagarajan, Andreas Spanias</td>
<td></td>
</tr>
<tr>
<td>ELLIPTICAL SHAPE RECOVERY FROM BLURRED PIXELS USING DEEP LEARNING</td>
<td>2775</td>
</tr>
<tr>
<td>Hojatollah Zamani, Peyman Rostami, Arash Amini, Farokh Marvasti</td>
<td></td>
</tr>
<tr>
<td>FACTORIZED CRF WITH BATCH NORMALIZATION BASED ON THE ENTIRE TRAINING DATA</td>
<td>2780</td>
</tr>
<tr>
<td>Eran Goldman, Jacob Goldberger</td>
<td></td>
</tr>
<tr>
<td><strong>MLSP-3: DEEP LEARNING TRAINING METHODS 3</strong></td>
<td></td>
</tr>
<tr>
<td>EVOLUTIONARY QUANTIZATION OF NEURAL NETWORKS WITH MIXED-PRECISION</td>
<td>2785</td>
</tr>
<tr>
<td>Zhenhua Liu, Xinfeng Zhang, Shanshe Wang, Siwei Ma, Wen Gao</td>
<td></td>
</tr>
<tr>
<td>EVOLVING QUANTIZED NEURAL NETWORKS FOR IMAGE CLASSIFICATION USING A MULTI-OBJECTIVE GENETIC ALGORITHM</td>
<td>2790</td>
</tr>
<tr>
<td>Yong Wang, Xiaojing Wang, Xiaoyu He</td>
<td></td>
</tr>
<tr>
<td>SPECTRAL DOMAIN CONVOLUTIONAL NEURAL NETWORK</td>
<td>2795</td>
</tr>
<tr>
<td>Bochen Guan, Jinnian Zhang, William A. Sethares, Richard Kijowski, Fang Liu</td>
<td></td>
</tr>
<tr>
<td>PARAMETRIC SPECTRAL FILTERS FOR FAST CONVERGING, SCALABLE CONVOLUTIONAL NEURAL NETWORKS</td>
<td>2800</td>
</tr>
</tbody>
</table>
FEATURE REUSE FOR A RANDOMIZATION BASED NEURAL NETWORK ........................................... 2805
   Xinyue Liang, Mikael Skoglund, Saikat Chatterjee

A RELU DENSE LAYER TO IMPROVE THE PERFORMANCE OF NEURAL NETWORKS .......... 2810
   Alireza M. Javid, Sandipan Das, Mikael Skoglund, Saikat Chatterjee

MLSP-4: MACHINE LEARNING FOR CLASSIFICATION APPLICATIONS 1

NESTED LEARNING FOR MULTI-LEVEL CLASSIFICATION ......................................................... 2815
   Raphaël Achddou, J. Matias Di Martino, Guillermo Sapiro

CROSS-MODAL REPRESENTATION RECONSTRUCTION FOR ZERO-SHOT
CLASSIFICATION ...................................................................................................................... 2820
   Yu Wang, Shenjie Zhao

HIGCNN: HIERARCHICAL INTERLEAVED GROUP CONVOLUTIONAL NEURAL
NETWORKS FOR POINT CLOUDS ANALYSIS .............................................................................. 2825
   Jisheng Dang, Jun Yang

AUTOKWS: KEYWORD SPOTTING WITH DIFFERENTIABLE ARCHITECTURE SEARCH .......... 2830
   Bo Zhang, Wenfeng Li, Qingyuan Li, Weiji Zhuang, Xiangxiang Chu, Yijun Wang

EMBEDDING SEMANTIC HIERARCHY IN DISCRETE OPTIMAL TRANSPORT FOR RISK
MINIMIZATION ......................................................................................................................... 2835
   Yubin Ge, Site Li, Xuyang Li, Fangfang Fan, Wanqing Xie, Jane You, Xiaofeng Liu

IDENTIFYING SPAMMERS TO BOOST CROWDSOURCED CLASSIFICATION ....................... 2840
   Panagiotis A. Traganitis, Georgios B. Giannakis

MLSP-5: MACHINE LEARNING FOR CLASSIFICATION APPLICATIONS 2

A RANK-CONSTRAINED CLUSTERING ALGORITHM WITH ADAPTIVE EMBEDDING .......... 2845
   Shenfei Pei, Feiping Nie, Rong Wang, Xuelong Li

TOWARDS EFFICIENT AGE ESTIMATION BY EMBEDDING POTENTIAL GENDER
FEATURES ......................................................................................................................................... 2850
   Yulan Deng, Lunke Fei, Shaohua Teng, Wei Zhang, Dongning Liu, Yan Hou

ADVERSARIAL ATTACKS ON COARSE-TO-FINE CLASSIFIERS ............................................... 2855
   Ismail R. Alkhouri, George K. Atia

GDTW: A NOVEL DIFFERENTIABLE DTW LOSS FOR TIME SERIES TASKS ....................... 2860
   Xiang Liu, Naiqi Li, Shu-Tao Xia

HIERARCHICAL RECURRENT NEURAL NETWORK FOR HANDWRITTEN STROKES
CLASSIFICATION ......................................................................................................................... 2865
   Illya Degtyarenko, Ivan Deriuga, Andrii Grygoriev, Serhii Polotskyi, Volodymyr Melnyk,
   Dmytro Zharkov, Olga Radyvonenko

ROBUST DOMAIN-FREE DOMAIN GENERALIZATION WITH CLASS-AWARE
ALIGNMENT ................................................................................................................................. 2870
   Wenyu Zhang, Mohamed Ragab, Ramon Sagarna
MLSP-6: COMPRESSED SENSING AND LEARNING

ONE-BIT COMPRESSED SENSING USING UNTRAINED NETWORK PRIOR .................................................. 2875
Swatantra Kafle, Geethu Joseph, Pramod K. Varshney

DEEP UNFOLDING NETWORK FOR BLOCK-SPARSE SIGNAL RECOVERY .............................................. 2880
Rong Fu, Vincent Monardo, Tianyao Huang, Yimin Liu

REST: ROBUST LEARNED SHRINKAGE-THRESHOLDING NETWORK TAMING INVERSE PROBLEMS WITH MODEL MISMATCH .................................................................................................. 2885
Wei Pu, Chao Zhou, Yonina C. Eldar, Miguel R. D. Rodrigues

UNFOLDING NEURAL NETWORKS FOR COMPRESSION MULTICHANNEL BLIND DECONVOLUTION .. 2890
Bahareh Tolooshams, Satish Mulleti, Demba Ba, Yonina C. Eldar

SPARSITY DRIVEN LATENT SPACE SAMPLING FOR GENERATIVE PRIOR BASED
COMPRESSIVE SENSING ........................................................................................................................... 2895
Vinayak Killedar, Praveen Kumar Pokala, Chandra Sekhar Seelamantula

A SPARSE CODING APPROACH TO AUTOMATIC DIET MONITORING WITH CONTINUOUS GLUCOSE MONITORS ........................................................................................................... 2900
Anurag Das, Seyedhooman Sajjadi, Bobak Mortazavi, Theodora Chaspari, Prajna Paromita, Laura Ruebush, Nicolaas Deutz, Ricardo Gutierrez-Osuna

MLSP-7: TENSOR SIGNAL PROCESSING

SPEEDING UP OF KERNEL-BASED LEARNING FOR HIGH-ORDER TENSORS ........................................ 2905
Ouafae Karmouda, Jeremie Boulanger, Remy Boyer

A FAST RANDOMIZED ADAPTIVE CP DECOMPOSITION FOR STREAMING TENSORS .......... 2910
Le Trung Thanh, Karim Abed-Meraim, Nguyen Link Trung, Adel Hafiane

RANK-REVEALING BLOCK-TERM DECOMPOSITION FOR TENSOR COMPLETION .............. 2915
Athanassios A. Rontogiannis, Paris V Giampouras, Eleftherios Kofidis

KERNEL LEARNING WITH TENSOR NETWORKS ....................................................................................... 2920
Kriton Konstantinidis, Shengxi Li, Danilo P. Mandic

FIBER-SAMPLED STOCHASTIC MIRROR DESCENT FOR TENSOR DECOMPOSITION
WITH β-DIVERGENCE ............................................................................................................................... 2925
Wenqiang Pu, Shahana Ibrahim, Xiao Fu, Mingyi Hong

REGULARIZED RECOVERY BY MULTI-ORDER PARTIAL HYPERGRAPH TOTAL
VARIATION .................................................................................................................................................. 2930
Ruyuan Qu, Jiaqi He, Hui Feng, Chongbin Xu, Bo Hu

MLSP-8: LEARNING

LEARNING DISCRIMINATIVE FEATURES FOR SEMI-SUPERVISED ANOMALY DETECTION .......................................................... 2935
Zhe Feng, Jie Tang, Yishun Dou, Gangshan Wu
RGLN: ROBUST RESIDUAL GRAPH LEARNING NETWORKS VIA SIMILARITY-PRESERVING MAPPING ON GRAPHS

Jiaxiang Tang, Xiang Gao, Wei Hu

SEQUENCE-LEVEL SELF-TEACHING REGULARIZATION

Eric Sun, Liang Lu, Zhong Meng, Yifan Gong

WEARING A MASK: COMPRESSED REPRESENTATIONS OF VARIABLE-LENGTH SEQUENCES USING RECURRENT NEURAL TANGENT KERNELS

Sina Alemohammad, Hossein Babaei, Randall Balestriero, Matt Y. Cheung, Ahmed Imtiaz Humayun, Daniel Lejeune, Naiming Liu, Lorenzo Luzi, Jasper Tan, Zichao Wang, Richard G. Baraniuk

H-GPR: A HYBRID STRATEGY FOR LARGE-SCALE GAUSSIAN PROCESS REGRESSION

Naiqi Li, Yinghua Gao, Wenjie Li, Yong Jiang, Shu-Tao Xia

LEARNING OPTIMAL LATTICE CODES FOR MIMO COMMUNICATIONS

Laia Amorós, Mikko Pitkänen

MLSP-9: LEARNING THEORY FOR NEURAL NETWORKS

A BAYESIAN INTERPRETATION OF THE LIGHT GATED RECURRENT UNIT

Alexandre Bittar, Philip N. Garner

A LARGE-DIMENSIONAL ANALYSIS OF SYMMETRIC SNE

Charles Séjourné, Romain Couillet, Pierre Comon

A DYNAMICAL SYSTEMS PERSPECTIVE ON ONLINE BAYESIAN NONPARAMETRIC ESTIMATORS WITH ADAPTIVE HYPERPARAMETERS

Alec Koppel, Amrit S. Bedi, Vikram Krishnamurthy

ONLINE MULTI-HOP INFORMATION BASED KERNEL LEARNING OVER GRAPHS

Zixiao Zong, Yanning Shen

SPARSITY IN MAX-PLUS ALGEBRA AND APPLICATIONS IN MULTIVARIATE CONVEX REGRESSION

Nikos Tsilivis, Anastasios Tsiamis, Petros Maragos

COMPLEX-VALUED VS. REAL-VALUED NEURAL NETWORKS FOR CLASSIFICATION PERSPECTIVES: AN EXAMPLE ON NON-CIRCULAR DATA

J. A. Barrassina, C. Ren, C. Morisseau, G. Vieillard, J.-P. Ovarlez

MLSP-10: DEEP LEARNING FOR SPEECH AND AUDIO

HIGH-FREQUENCY ADVERSARIAL DEFENSE FOR SPEECH AND AUDIO

R. Olivier, B. Raj, M. Shah

LEARNING SEPARABLE TIME-FREQUENCY FILTERBANKS FOR AUDIO CLASSIFICATION

Jie Pu, Yannis Panagakis, Maja Pantic

UPSAMPLING ARTIFACTS IN NEURAL AUDIO SYNTHESIS

Jordi Pons, Santiago Pascual, Giulio Cengarle, Joan Serrà
DEEP CONVOLUTIONAL AND RECURRENT NETWORKS FOR POLYPHONIC INSTRUMENT CLASSIFICATION FROM MONOPHONIC RAW AUDIO WAVEFORMS ........................................... 3010
Kleenthis Avramidis, Agelos Kratimenos, Christos Garoufis, Athanasia Zlatintsi, Petros Maragos

LEARNING AUDIO EMBEDDINGS WITH USER LISTENING DATA FOR CONTENT-BASED MUSIC RECOMMENDATION ...................................................................................................... 3015
Ke Chen, Beici Liang, Xiaoshuan Ma, Minwei Gu

EFFICIENT SPEECH EMOTION RECOGNITION USING MULTI-SCALE CNN AND ATTENTION .......................................................................................................................... 3020
Zixuan Peng, Yu Lu, Shengfeng Pan, Yunfeng Liu

MLSP-11: SELF-SUPERVISED LEARNING FOR SPEECH PROCESSING

NEURAL AUDIO FINGERPRINT FOR HIGH-SPECIFIC AUDIO RETRIEVAL BASED ON CONTRASTIVE LEARNING ........................................................................................................... 3025
Sungkyun Chang, Donmoon Lee, Jeongsoo Park, Hyungui Lim, Kyogu Lee, Karam Ko, Yoonchang Han

SELF-TRAINING AND PRE-TRAINING ARE COMPLEMENTARY FOR SPEECH RECOGNITION .......................................................................................................................... 3030
Qiantong Xu, Alexei Baevski, Tatiana Likhomanenko, Paden Tomasello, Alexis Conneau, Ronan Collobert, Gabriel Synnaeve, Michael Auli

UNSUPERVISED DISCRIMINATIVE LEARNING OF SOUNDS FOR AUDIO EVENT CLASSIFICATION ................................................................................................................. 3035
Sascha Hornauer, Ke Li, Stella X. Yu, Shabnam Ghaffarzadegan, Liu Ren

SIMILARITY ANALYSIS OF SELF-SUPERVISED SPEECH REPRESENTATIONS .......................................................................................................................... 3040
Yu-An Chung, Yonatan Belinkov, James Glass

JOINT Masked CPC AND CTC TRAINING FOR ASR .................................................................................. 3045
Chaitanya Talnikar, Tatiana Likhomanenko, Ronan Collobert, Gabriel Synnaeve

A COMPARISON OF DISCRETE LATENT VARIABLE MODELS FOR SPEECH REPRESENTATION LEARNING .......................................................................................... 3050
Henry Zhou, Alexei Baevski, Michael Auli

MLSP-12: FEDERATED LEARNING 1

FEDERATED LEARNING FROM BIG DATA OVER NETWORKS ........................................................................... 3055
Y. Sarcheshmehpour, M. Leinonen, A. Jung

EFFICIENT CLIENT CONTRIBUTION EVALUATION FOR HORIZONTAL FEDERATED LEARNING .................................................................................................................. 3060
Jie Zhao, Xinghua Zhu, Jianzong Wang, Jing Xiao

A QUANTITATIVE METRIC FOR PRIVACY LEAKAGE IN FEDERATED LEARNING ............................................................................ 3065
Yong Liu, Xinghua Zhu, Jianzong Wang, Jing Xiao

DP-SIGNSGD: WHEN EFFICIENCY MEETS PRIVACY AND ROBUSTNESS ............................................................................ 3070
Lingjuan Lyu
FEDERATED ALGORITHM WITH BAYESIAN APPROACH: OMNI-FEDGE .................................................. 3075
Sai Anuroop Kesapalli, B. N. Bharath

TRAINING SPEECH RECOGNITION MODELS WITH FEDERATED LEARNING: A QUALITY/COST FRAMEWORK ................................................................................................................ 3080
Dhruv Guliani, Françoise Beaufays, Giovanni Motta

MLSP-13: FEDERATED LEARNING 2

CROSS-SILO FEDERATED TRAINING IN THE CLOUD WITH DIVERSITY SCALING AND SEMI-SUPERVISED LEARNING................................................................................................................ 3085
Kishore Nandury, Anand Mohan, Frederick Weber

GRADUAL FEDERATED LEARNING USING SIMULATED ANNEALING........................................... 3090
Luong Trung Nguyen, Byonghyo Shim

OPTIMAL IMPORTANCE SAMPLING FOR FEDERATED LEARNING.................................................. 3095
Elsa Rizk, Stefan Vlaski, Ali H. Sayed

MULTI-TIER FEDERATED LEARNING FOR VERTICALLY PARTITIONED DATA................................. 3100
Anirban Das, Stacy Patterson

ENERGY MINIMIZATION FOR FEDERATED LEARNING WITH IRS-ASSISTED OVER-THE-AIR COMPUTATION.......................................................................................................................... 3105
Yuntao Hu, Ming Chen, Mingche Chen, Zhaohui Yang, Mohammad Shikh-Bahaei, H. Vincent Poor, Shuguang Cui

ADAPTIVE QUANTIZATION OF MODEL UPDATES FOR COMMUNICATION-EFFICIENT FEDERATED LEARNING .......................................................................................................................... 3110
Divyansh Jhunjhunwala, Advait Gadhikar, Gauri Joshi, Yonina C. Eldar

MLSP-14: LEARNING ALGORITHMS 1

HEBBNET: A SIMPLIFIED HEBBIAN LEARNING FRAMEWORK TO DO BIOLOGICALLY PLAUSIBLE LEARNING.................................................................................................................. 3115
Manas Gupta, Arulmurugan Ambikapathi, Savitha Ramasamy

T-K-MEANS: A ROBUST AND STABLE K-MEANS VARIANT ............................................................. 3120
Yiming Li, Yang Zhang, Qingtao Tang, Weipeng Huang, Yong Jiang, Shu-Tao Xia

ADAPTIVE FEATURE WEIGHT LEARNING FOR ROBUST CLUSTERING PROBLEM WITH SPARSE CONSTRAINT .......................................................................................................................... 3125
Feiping Nie, Wei Chang, Xuelong Li, Jin Xu, Gongfu Li

ASSISTED LEARNING: COOPERATIVE AI WITH AUTONOMY ......................................................... 3130
Jiaying Zhou, Xun Xian, Na Li, Jie Ding

GEOM-SPIDER-EM: FASTER VARIANCE REDUCED STOCHASTIC EXPECTATION MAXIMIZATION FOR NONCONVEX FINITE-SUM OPTIMIZATION .................................................. 3135
Gersende Fort, Eric Moulines, Hoi-To Wai

LEARNING A TREE OF NEURAL NETS ............................................................................................... 3140
Arman Zharmagambetov, Miguel A. Carreira-Perpiñán
MLSP-15: LEARNING ALGORITHMS 2

CORRUPTED CONTEXTUAL BANDITS: ONLINE LEARNING WITH CORRUPTED CONTEXT .......................................................... 3145
  Djallel Bouneffouf

TRAINING A BANK OF WIENER MODELS WITH A NOVEL QUADRATIC MUTUAL INFORMATION COST FUNCTION ........................................ 3150
  Bo Hu, Jose C. Principe

INFORMATION AND REGULARIZATION IN RESTRICTED BOLTZMANN MACHINES .......................................................... 3155
  Matias Vera, Leonardo Rey Vega, Pablo Piantanida

DEEP DETERMINISTIC INFORMATION BOTTLENECK WITH MATRIX-BASED ENTROPY FUNCTION .......................................................... 3160
  Xi Yu, Shujian Yu, José C. Príncipe

TRANSITIVE TRANSFER SPARSE CODING FOR DISTANT DOMAIN .................................................................................. 3165
  Lingtian Feng, Feng Qian, Xin He, Yuqi Fan, Hanpeng Cai, Guangmin Hu

FAST LOCAL REPRESENTATION LEARNING WITH ADAPTIVE ANCHOR GRAPH .......................................................... 3170
  Canyu Zhang, Feiping Nie, Zheng Wang, Rong Wang, Xuelong Li

MLSP-16: ML AND GRAPHS

LEARNING ON HETEROGENEOUS GRAPHS USING HIGH-ORDER RELATIONS .................................................................. 3175
  See Hian Lee, Feng Ji, Wee Peng Tay

INCOMPLETE MULTI-VIEW SUBSPACE CLUSTERING WITH LOW-RANK TENSOR .......................................................... 3180
  Jianlun Liu, Shaohua Teng, Wei Zhang, Xiaozhao Fang, Lunke Fei, Zhuxiu Zhang

CHANNEL-WISE MIX-FUSION DEEP NEURAL NETWORKS FOR ZERO-SHOT LEARNING .................................................. 3185
  Guowei Wang, Naiyang Guau, Hanjia Ye, Xiaodong Yi, Hang Cheng, Junjie Zhu

ONLINE UNSUPERVISED LEARNING USING ENSEMBLE GAUSSIAN PROCESSES WITH RANDOM FEATURES ................. 3190
  Georgios V Karanikolas, Qin Lu, Georgios B. Giannakis

DIMENSION SELECTED SUBSPACE CLUSTERING .......................................................................................................................... 3195
  Shuoyang Li, Yuhui Luo, Jonathon Chambers, Wenwu Wang

DEEP ENSEMBLE SIAMESE NETWORK FOR INCREMENTAL SIGNAL CLASSIFICATION ......................................................... 3200
  Chen Yang, Shuyuan Yang

MLSP-17: GRAPH NEURAL NETWORKS

NON-RECURSIVE GRAPH CONVOLUTIONAL NETWORKS .................................................................................................................. 3205
  Hao Chen, Zengde Deng, Yue Xu, Zhoujun Li

EGO-BASED ENTROPY MEASURES FOR STRUCTURAL REPRESENTATIONS ON GRAPHS .......................................................... 3210
  George Dasoulas, Giannis Nikolentzos, Kevin Seaman, Aladin Virmaux, Michalis Vazirgiannis
SYMmetric Sub-Graph Spatio-Temporal Graph Convolution and Its Application in Complex Activity Recognition .......................................................... 3215
Pratyusha Das, Antonio Ortega

Progressive Spatio-Temporal Graph Convolutional Network for Skeleton-Based Human Action Recognition .......................................................... 3220
Negar Heidari, Alexandros Iosifidis

Sparse-Coded Dynamic Mode Decomposition on Graph for Prediction of River Water Level Distribution .......................................................... 3225
Yusuke Arai, Shogo Muramatsu, Hiroyasu Yasuda, Kiyoshi Hayasaka, Yu Otake

Graph Frequency Analysis of COVID-19 Incidence to Identify County-Level Contagion Patterns in the United States .................................................. 3230
Yang Li, Gonzalo Mateos

MLSP-18: Matrix Factorization and Applications

Generalized Polytopic Matrix Factorization .......................................................... 3235
Gokcan Tatli, Alper T. Erdogan

Exact Linear Convergence Rate Analysis for Low-Rank Symmetric Matrix Completion via Gradient Descent ......................................................... 3240
Trung Vu, Raviv Raich

Structured Support Exploration for Multilayer Sparse Matrix Factorization .......................................................... 3245
Quoc-Tung Le, Rémi Gribonval

Optimal Selection of Matrix Shape and Decomposition Scheme for Neural Network Compression .......................................................... 3250
Yerlan Idelbayev, Miguel A. Carreira-Perpiñan

Sparse Graph Based Sketching for Fast Numerical Linear Algebra .......................................................... 3255
Dong Hu, Shashanka Ubaru, Alex Gittens, Kenneth L. Clarkson, Lior Horesh, Vassilis Kalantzis

Cold Start Revisited: A Deep Hybrid Recommender with Cold-Warm Item Harmonization .......................................................... 3260
Oren Barkan, Roy Hirsch, Ori Katz, Avi Caciularu, Yoni Weill, Noam Koenigstein

MLSP-19: Non-Negative Matrix Factorization

On a Guided Nonnegative Matrix Factorization .......................................................... 3265
Joshua Vendrow, Jamie Haddock, Elizaveta Rebrova, Deanna Needell

Nonnegative Unimodal Matrix Factorization .......................................................... 3270
Andersen Man Shun Ang, Nicolas Gillis, Arnaud Vandaele, Hans De Sterck

Kernel Orthogonal Nonnegative Matrix Factorization: Application to Multispectral Document Image Decomposition .................................................. 3275
Abderrahmane Rahiche, Mohamed Cheriet
RANDOM PROJECTION STREAMS FOR (WEIGHTED) NONNEGATIVE MATRIX FACTORIZATION ........................................................................................................................................ 3280
  Farouk Yahaya, Matthieu Puigt, Gilles Delmaire, Gilles Roussel

MULTIVARIATE NON-NEGATIVE MATRIX FACTORIZATION WITH APPLICATION TO ENERGY DISAGGREGATION ................................................................. 3285
  Pascal A. Schirmer, Iosif Mporas

MLSP-20: ATTENTION AND AUTOENCODER NETWORKS

CONTINUOUS-TIME SELF-ATTENTION IN NEURAL DIFFERENTIAL EQUATION ................................................................. 3290
  Jen-Tzung Chien, Yi-Hsiang Chen

BLIND DEINTERLEAVING OF SIGNALS IN TIME SERIES WITH SELF-ATTENTION BASED SOFT MIN-COST FLOW LEARNING ................................................................................................................... 3295
  Ogul Can, Yeti Z. Gurbuz, Berkin Yildirim, A. Aydin Alatan

ATTENTION ON ATTENTION SPARSE DENSE CONVOLUTIONAL NETWORK FOR FINANCIAL SIGNAL PROCESSING .................................................................................................................. 3300
  Tianlei Zhu, Jiawei Li, Xinji Liu, Yong Jiang, Shu-Tao Xia

STOCK MOVEMENT PREDICTION AND PORTFOLIO MANAGEMENT VIA MULTIMODAL LEARNING WITH TRANSFORMER .......................................................................................................... 3305
  Divyanshu Daiya, Che Lin

A QUATERNION-VALUED VARIATIONAL AUTOENCODER ......................................................................................................................... 3310
  Eleonora Grassucci, Danilo Comminiello, Aurelio Uncini

LEARNING A SPARSE GENERATIVE NON-PARAMETRIC SUPERVISED AUTOENCODER.................................................................................................................. 3315
  Michel Barlaud, Frederic Guyard

MLSP-21: GENERATIVE NEURAL NETWORKS

DAG-GAN: CAUSAL STRUCTURE LEARNING WITH GENERATIVE ADVERSARIAL NETS .............................................................................................................. 3320
  Yinghua Gao, Li Shen, Shu-Tao Xia

RELAXED WASSERSTEIN WITH APPLICATIONS TO GANS .............................................................................................................................. 3325
  Xin Guo, Johnny Hong, Tianyi Lin, Nan Yang

ENVIRONMENT-INDEPENDENT WI-FI HUMAN ACTIVITY RECOGNITION WITH ADVERSARIAL NETWORK .......................................................................................................................... 3330
  Zhengyang Wang, Sheng Chen, Wei Yang, Yang Xu

A ROBUST TO NOISE ADVERSARIAL RECURRENT MODEL FOR NON-INTRUSIVE LOAD MONITORING ................................................................................................. 3335
  Maria Kaselini, Athanasios Voulodimos, Nikolaos Doulamis, Anastasios Doulamis, Efthychios Protopapadakis

ENHANCING DATA-FREE ADVERSARIAL DISTILLATION WITH ACTIVATION REGULARIZATION AND VIRTUAL INTERPOLATION ........................................................................................................... 3340
  Xiaoyang Qu, Jianzong Wang, Jing Xiao

SEQUENTIAL ADVERSARIAL ANOMALY DETECTION WITH DEEP FOURIER KERNEL ...................................................................................................... 3345
  Shixiang Zhu, Henry Shaowu Yuchi, Minghe Zhang, Yao Xie
MLSP-22: SEQUENTIAL LEARNING

INCORPORATE MAXIMUM MEAN DISCREPANCY IN RECURRENT LATENT SPACE FOR SEQUENTIAL GENERATIVE MODEL ................................................................. 3350
Yuchi Zhang, Yongliang Wang, Yang Dong

FMA-ETA: ESTIMATING TRAVEL TIME ENTIRELY BASED ON FFN WITH ATTENTION .......... 3355
Yiwen Sun, Yulu Wang, Kun Fu, Zheng Wang, Ziang Yan, Changshui Zhang, Jieping Ye

A UNIFIED APPROACH TO TRANSLATE CLASSICAL BANDIT ALGORITHMS TO STRUCTURED BANDITS .............................................................. 3360
Samarth Gupta, Shreyas Chaudhari, Subhojyoti Mukherjee, Gauri Joshi, Osman Yagan

NEAR-OPTIMAL ALGORITHMS FOR PIECEWISE-STATIONARY CASCADING BANDITS ........ 3365
Lingda Wang, Huozi Zhou, Bingcong Li, Lav R. Varshney, Zhizhen Zhao

OPTIMUM FEATURE ORDERING FOR DYNAMIC INSTANCE–WISE JOINT FEATURE SELECTION AND CLASSIFICATION ....................................................... 3370
Yasitha Warahena Liyanage, Daphney-Stavroula Zois

POLA: ONLINE TIME SERIES PREDICTION BY ADAPTIVE LEARNING RATES .................. 3375
Wenyu Zhang

MLSP-23: APPLICATIONS IN MUSIC AND AUDIO PROCESSING

SINGER IDENTIFICATION USING DEEP TIMBRE FEATURE LEARNING WITH KNN-NET ........ 3380
Xulong Zhang, Jiale Qian, Yi Yu, Yifu Sun, Wei Li

IMPLICIT HRTF MODELING USING TEMPORAL CONVOLUTIONAL NETWORKS ............ 3385
Israel D. Gebru, Dejan Markovic, Alexander Richard, Steven Krenn, Gladstone A. Butler, Fernando De La Torre, Yaser Sheikh

IMPROVING THE CLASSIFICATION OF RARE CHORDS WITH UNLABELED DATA .......... 3390
Marcelo Bortolazzo, Rodrigo Schramm, Claudio R. Jung

LOOPNET: MUSICAL LOOP SYNTHESIS CONDITIONED ON INTUITIVE MUSICAL PARAMETERS ............................................................... 3395
Pritish Chandna, António Ramires, Xavier Serra, Emilia Gómez

MICAUGMENT: ONE-SHOT MICROPHONE STYLE TRANSFER .................................. 3400
Zalán Borsos, Yunpeng Li, Beat Gfeller, Marco Tagliasacchi

MLSP-24: APPLICATIONS IN AUDIO AND SPEECH PROCESSING

WASSERSTEIN BARYCENTER TRANSPORT FOR ACOUSTIC ADAPTATION .................. 3405
Eduardo F. Montesuma, Fred-Maurice Ngolè Mboula

EFFICIENT ADVERSARIAL AUDIO SYNTHESIS VIA PROGRESSIVE UPSAMPLING .......... 3410
Youngwoo Cho, Minwook Chang, Sanghyeon Lee, Hyoungwoo Lee, Gerard Jounghyun Kim, Jaegul Choo

MULTI-CHANNEL SPEECH ENHANCEMENT USING GRAPH NEURAL NETWORKS ............ 3415
Panagiotis Tzirakis, Anurag Kumar, Jacob Donley
<table>
<thead>
<tr>
<th>Title</th>
<th>Authors</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>MULTI-DECODER DPRNN: SOURCE SEPARATION FOR VARIABLE NUMBER OF SPEAKERS</td>
<td>Junzhe Zhu, Raymond A. Yeh, Mark Hasegawa-Johnson</td>
<td>3420</td>
</tr>
<tr>
<td>DATA-EFFICIENT FRAMEWORK FOR REAL-WORLD MULTIPLE SOUND SOURCE 2D LOCALIZATION</td>
<td>Guillaume Le Moing, Phongtharin Vinayavekhin, Don Joven Agravanite, Tadanobu Inoue, Jayakorn Yongkulbhisal, Asim Munawar, Ryuki Tachibana</td>
<td>3425</td>
</tr>
<tr>
<td>FUSING INFORMATION STREAMS IN END-TO-END AUDIO-VISUAL SPEECH RECOGNITION</td>
<td>Wentao Yu, Steffen Zeiler, Dorothea Kolossa</td>
<td>3430</td>
</tr>
<tr>
<td><strong>MLSP-25: REINFORCEMENT LEARNING 1</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COOPERATIVE SCENARIOS FOR MULTI-AGENT REINFORCEMENT LEARNING IN WIRELESS EDGE CACHING</td>
<td>Navneet Garg, Tharmalingam Ratnarajah</td>
<td>3435</td>
</tr>
<tr>
<td>ROBUST DEEP REINFORCEMENT LEARNING FOR UNDERWATER NAVIGATION WITH UNKNOWN DISTURBANCES</td>
<td>Juan Parras, Santiago Zazo</td>
<td>3440</td>
</tr>
<tr>
<td>ONLINE HYPER-PARAMETER TUNING FOR THE CONTEXTUAL BANDIT</td>
<td>Djallel Bouneffouf, Emmanuelle Claey</td>
<td>3445</td>
</tr>
<tr>
<td>DOUBLE-LINEAR THOMPSON SAMPLING FOR CONTEXT-ATTENTIVE BANDITS</td>
<td>Djallel Bouneffouf, Raphael Feraud, Sohini Upadhyay, Yasaman Khazaeni, Irina Rish</td>
<td>3450</td>
</tr>
<tr>
<td>ON THE MARGINAL BENEFIT OF ACTIVE LEARNING: DOES SELF-SUPERVISION EAT ITS CAKE?</td>
<td>Yao-Chun Chan, Mingchen Li, Samet Oymak</td>
<td>3455</td>
</tr>
<tr>
<td>ROBUST MAML: PRIORITIZATION TASK BUFFER WITH ADAPTIVE LEARNING PROCESS FOR MODEL-AGNOSTIC META-LEARNING</td>
<td>Thanh Nguyen, Tung Luu, Trung Pham, Sanzhar Rakhimkul, Chang D. Yoo</td>
<td>3460</td>
</tr>
<tr>
<td><strong>MLSP-26: REINFORCEMENT LEARNING 2</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>INTRODUCING DEEP REINFORCEMENT LEARNING TO NLU RANKING TASKS</td>
<td>Ge Yu, Emre Barut, Chengwei Su</td>
<td>3465</td>
</tr>
<tr>
<td>TEMPORAL LINK PREDICTION VIA REINFORCEMENT LEARNING</td>
<td>Ye Tao, Ying Li, Zhonghai Wu</td>
<td>3470</td>
</tr>
<tr>
<td>A DEEP REINFORCEMENT LEARNING APPROACH TO AUDIO-BASED NAVIGATION IN A MULTI-SPEAKER ENVIRONMENT</td>
<td>Petros Giannakopoulos, Aggelos Pikrakis, Yannis Cotronis</td>
<td>3475</td>
</tr>
<tr>
<td>GLOBAL-LOCALIZED AGENT GRAPH CONVOLUTION FOR MULTI-AgENT REINFORCEMENT</td>
<td>Yuntao Liu, Yong Dou, Siqi Shen, Peng Qiao</td>
<td>3480</td>
</tr>
</tbody>
</table>
MLSP-27: REINFORCEMENT LEARNING 3

GAUSSIAN PROCESS TEMPORAL-DIFFERENCE LEARNING WITH SCALABILITY AND WORST-CASE PERFORMANCE GUARANTEES................................................................. 3485
Qin Lu, Georgios B. Giannakis

SELF-INFERENCE OF OTHERS’ POLICIES FOR HOMOGENEOUS AGENTS IN COOPERATIVE MULTI-AGENT REINFORCEMENT LEARNING .................................................. 3490
Qifeng Lin, Qing Ling

SEMI-SUPERVISED BATCH ACTIVE LEARNING VIA BILEVEL OPTIMIZATION........................ 3495
Zalán Borsos, Marco Tagliasacchi, Andreas Krause

KERNEARL-BASED LIFELONG POLICY GRADIENT REINFORCEMENT LEARNING........ 3500
Rami Mowakeaa, Seung-Jun Kim, Darren K. Emge

POLICY AUGMENTATION: AN EXPLORATION STRATEGY FOR FASTER CONVERGENCE OF DEEP REINFORCEMENT LEARNING ALGORITHMS.............................. 3505
Arash Mahyari

GRAPHCOMM: A GRAPH NEURAL NETWORK BASED METHOD FOR MULTI-AGENT REINFORCEMENT LEARNING .......................................................... 3510
Siqi Shen, Yongquan Fu, Huayou Su, Hengyue Pan, Peng Qiao, Yong Dou, Cheng Wang

MLSP-28: ML AND TIME SERIES

IN SITU CALIBRATION OF CROSS-SENSITIVE SENSORS IN MOBILE SENSOR ARRAYS USING FAST INFORMED NON-NEGATIVE MATRIX FACTORIZATION................................... 3515
Olivier Vu Thanh, Matthieu Puigt, Farouk Yahaya, Gilles Delmaire, Gilles Roussel

MULTIPHISH: MULTI-MODAL FEATURES FUSION NETWORKS FOR PHISHING DETECTION................................................................. 3520
Lei Zhang, Peng Zhang, Luchen Liu, Jianlong Tan

FAILURE PREDICTION BY CONFIDENCE ESTIMATION OF UNCERTAINTY-AWARE DIRICHLET NETWORKS......................................................... 3525
Theodoros Tsiligkaridis

TWO-STAGE FRAMEWORK FOR SEASONAL TIME SERIES FORECASTING ....................... 3530
Qingyang Xu, Qingsong Wen, Liang Sun

RECURSIVE INPUT AND STATE ESTIMATION: A GENERAL FRAMEWORK FOR LEARNING FROM TIME SERIES WITH MISSING DATA........................................... 3535
Alberto García-Durán, Robert West

ON THE PERFORMANCE-COMPLEXITY TRADEOFF IN STOCHASTIC GREEDY WEAK SUBMODULAR OPTIMIZATION......................................................... 3540
Abolfazl Hashemi, Haris Vikalo, Gustavo De Veciana

MLSP-29: DEEP LEARNING FOR TIME SERIES

SEMI-SUPERVISED TIME SERIES CLASSIFICATION BY TEMPORAL RELATION PREDICTION ........................................................................................................... 3545
Haoyi Fan, Fengbin Zhang, Ruidong Wang, Xunhua Huang, Zuoyong Li
CONTINUOUS CNN FOR NONUNIFORM TIME SERIES .......................................................... 3550
Hui Shi, Yang Zhang, Hao Wu, Shiyou Chang, Kaizhi Qian, Mark Hasegawa-Johnson, Jishen Zhao

BLEND-RES\textsuperscript{2}NET: BLENDED REPRESENTATION SPACE BY TRANSFORMATION OF RESIDUAL MAPPING WITH RESTRAINED LEARNING FOR TIME SERIES CLASSIFICATION .......................................................... 3555
Arijit Ukil, Antonio J. Jara, Leandro Marin

SPATIOTEMPORAL ATTENTION FOR MULTIVARIATE TIME SERIES PREDICTION AND INTERPRETATION ........................................................................................................................................ 3560
Tryambak Gangopadhyay, Sin Yong Tan, Zhanhong Jiang, Rui Meng, Soumik Sarkar

TABULAR TRANSFORMERS FOR MODELING MULTIVARIATE TIME SERIES .................. 3565
Inkit Padhi, Yair Schiff, Igor Melnyk, Mattia Rigotti, Youssef Mroueh, Pierre Dognin, Jerret Ross, Ravi Nair, Erik Altman

REAL-TIME SYNCHRONIZATION IN NEURAL NETWORKS FOR MULTIVARIATE TIME SERIES ANOMALY DETECTION .............................................................................................................. 3570
Ahmed Abdulaal, Tomer Lancewicki

MLSP-30: GRAPH SIGNAL PROCESSING

FAST GRAPH KERNEL WITH OPTICAL RANDOM FEATURES .................................................. 3575
Hashem Ghanem, Nicolas Keriven, Nicolas Tremblay

FAST HIERARCHY PRESERVING GRAPH EMBEDDING VIA SUBSPACE CONSTRAINTS ........ 3580
Xu Chen, Lun Du, Mengyuan Chen, Yun Wang, Qingqing Long, Kunqing Xie

GRAPH EMBEDDING USING MULTI-LAYER ADJACENT POINT MERGING MODEL .......... 3585
Jianming Huang, Hiroyuki Kasai

NODE ATTRIBUTE COMPLETION IN KNOWLEDGE GRAPHS WITH MULTI-RELATIONAL PROPAGATION .......................................................... 3590
Eda Bayram, Alberto García-Durán, Robert West

MLSP-31: RECOMMENDATION SYSTEMS

USERREG: A SIMPLE BUT STRONG MODEL FOR RATING PREDICTION ......................... 3595
Haiyang Zhang, Ivan Ganchev, Nikola S. Nikolov, Mark Stevenson

TOWARD SKILLS DIALOG ORCHESTRATION WITH ONLINE LEARNING .................. 3600
Djallel Bouneffouf, Raphael Feraud, Sohini Upadhyay, Mayank Agarwal, Yasaman Khazaeni, Irina Rish

ADAPTIVE RE-BALANCING NETWORK WITH GATE MECHANISM FOR LONG-TAILED VISUAL QUESTION ANSWERING .......................................................... 3605
Hongyu Chen, Ruifang Liu, Han Fang, Ximing Zhang

CO-CAPSULE NETWORKS BASED KNOWLEDGE TRANSFER FOR CROSS-DOMAIN RECOMMENDATION .......................................................... 3610
Huiyuan Li, Li Yu, Youfang Leng, Qihan Du

MODUREC: RECOMMENDER SYSTEMS WITH FEATURE AND TIME MODULATION .......... 3615
Javier Maroto, Clément Vignac, Pascal Frossard
SIG2SIG: SIGNAL TRANSLATION NETWORKS TO TAKE THE REMAINS OF THE PAST............ 3620
  Sangyeon Kim, Hyunwoo Lee, Jonghee Han, Joon-Ho Kim

MLSP-32: OPTIMIZATION ALGORITHMS FOR MACHINE LEARNING

SOLVING A CLASS OF NON-CONVEX MIN-MAX GAMES USING ADAPTIVE MOMENTUM METHODS........................................................................................................................................................................ 3625
  Babak Barazandeh, Davoud Ataei Tarzanagh, George Michailidis

MINIMIZING WEIGHTED CONCAVE IMPURITY PARTITION UNDER CONSTRAINTS ........ 3630
  Thuan Nguyen, Think Nguyen

CONSTANT APPROXIMATION ALGORITHM FOR MINIMIZING CONCAVE IMPURITY........ 3635
  Thuan Nguyen, Hoang Le, Thinh Nguyen

FUSING MULTITASK MODELS BY RECURSIVE LEAST SQUARES ........................................ 3640
  Xiaobin Li, Lianlei Shan, Weiqiang Wang

CENTRALITY BASED NUMBER OF CLUSTER ESTIMATION IN GRAPH CLUSTERING......... 3645
  Mahdi Shamsi, Soosan Beheshti

DEPENDENCE-GUIDED MULTI-VIEW CLUSTERING ................................................................. 3650
  Xia Dong, Danyang Wu, Feiping Nie, Rong Wang, Xuelong Li

MLSP-33: OPTIMIZATION METHODS

IMPROVED STEP-SIZE SCHEDULES FOR NOISY GRADIENT METHODS............................ 3655
  Sarit Khirirat, Xiaoyu Wang, Sindri Magnússon, Mikael Johansson

RESPipe: RESILIENT MODEL-DISTRIBUTED DNN TRAINING AT EDGE NETWORKS ........ 3660
  Pengzhen Li, Erdem Koyuncu, Hulya Seferoglu

AN OPTIMAL STOCHASTIC COMPOSITIONAL OPTIMIZATION METHOD WITH APPLICATIONS TO META LEARNING................................................................. 3665
  Yuejiao Sun, Tianyi Chen, Wotao Yin

DECENTRALIZED OPTIMIZATION ON TIME-VARYING DIRECTED GRAPHS UNDER COMMUNICATION CONSTRAINTS......................................................... 3670
  Yiyue Chen, Abolfazl Hashemi, Haris Vikalo

DECENTRALIZED DEEP LEARNING USING MOMENTUM-ACCELERATED CONSENSUS .... 3675
  Aditya Balu, Zhanhong Jiang, Sin Yong Tan, Chinmay Hedge, Young M Lee, Soumik Sarkar

DEMystifying MODEL Averaging FOR COMMUNICATION-Efficient FEDERATED MATRIX FACTORIZATION .................................................................................. 3680
  Shuai Wang, Richard Cornelius Sawandi, Tsung-Hui Chang

MLSP-34: SUBSPACE LEARNING AND APPLICATIONS

SAMPLE EFFICIENT SUBSPACE-BASED REPRESENTATIONS FOR NONLINEAR META-LEARNING................................................................. 3685
  Halil Ibrahim Gulluk, Yue Sun, Samet Oymak, Maryam Fazel
MULTI-TASK LEARNING VIA SHARING INEXACT LOW-RANK SUBSPACE ........................................... Xiaoqian Wang, Feiping Nie

ON THE ADVERSARIAL ROBUSTNESS OF PRINCIPAL COMPONENT ANALYSIS .......................... Ying Li, Fuwei Li, Lifeng Lai, Jun Wu

FAST MANIFOLD LANDMARKING USING EXTREME EIGEN-PAIRS ............................................ Fen Wang, Gene Cheung, Yongchao Wang, Wai-Tian Tan

AFFINE PROJECTION SUBSPACE TRACKING ................................................................................. Marc Vilà, Carlos Alejandro López, Jaume Riba

A HIERARCHICAL SUBSPACE MODEL FOR LANGUAGE-ATTUNED ACOUSTIC UNIT DISCOVERY ................................................................................................................................................. Bolaji Yusuf, Lucas Ondel, Lukáš Burget, Jan Cernocký, Murat Saraçlar

MLSP-35: INDEPENDENT COMPONENT ANALYSIS

INDEPENDENT VECTOR ANALYSIS USING SEMI-PARAMETRIC DENSITY ESTIMATION VIA MULTIVARIATE ENTROPY MAXIMIZATION ......................................................................................................................... 3715 Lucas P. Damasceno, Charles C. Cavalcante, Tuğay Adali, Zois Boukouvalas

ICA WITH ORTHOGONALITY CONSTRAINT: IDENTIFIABILITY AND A NEW EFFICIENT ALGORITHM ............................................................................................................................................................................. 3720 Ben Gabrielson, M. A. B. S. Akhonda, Zois Boukouvalas, Seung-Jun Kim, Tuğay Adali

BLIND EXTRACTION OF MOVING SOURCES VIA INDEPENDENT COMPONENT AND VECTOR ANALYSIS: EXAMPLES ............................................................................................................................................................................. 3725 N. Amor, J. Cmejla, V. Kautský, Z. Koldovský, T. Kounovský

SINGLE CHANNEL VOICE SEPARATION FOR UNKNOWN NUMBER OF SPEAKERS UNDER REVERBERANT AND NOISY SETTINGS ............................................................................................................................................................................. 3730 Shlomo E. Chazan, Lior Wolf, Eliya Nachmani, Yossi Adi

UNSUPERVISED MUSICAL TIMBRE TRANSFER FOR NOTIFICATION SOUNDS .......................... Jing Yang, Tristan Cinquin, Gábor Sörös

MLSP-36: PATTERN RECOGNITION AND CLASSIFICATION

VISUAL PRIVACY PROTECTION VIA MAPPING DISTORTION .......................................................... 3740 Yiming Li, Peidong Liu, Yong Jiang, Shu-Tao Xia

L-RED: EFFICIENT POST-TRAINING DETECTION OF IMPERCEPTIBLE BACKDOOR ATTACKS WITHOUT ACCESS TO THE TRAINING SET ............................................................................................................. 3745 Zhen Xiang, David J. Miller, George Kesidis

MULTI-VIEW CONTRASTIVE LEARNING FOR ONLINE KNOWLEDGE DISTILLATION ............. 3750 Chuanguang Yang, Zhulin An, Yongqian Xu

DYNAMIC TEXTURE RECOGNITION VIA NUCLEAR DISTANCES ON KERNELIZED SCATTERING HISTOGRAM SPACES ............................................................................................................................................................................. 3755 Alexander Sagel, Julian Wörmann, Hao Shen
CLUSTERING A COLLECTION OF NETWORKS WITH MIXTURES OF L1-SPARSE GRAPHICAL MODELS

Zuogong Yue, Victor Solo .............................................................................................................. 3760

TAKING A CLOSER LOOK AT SYNTHESIS: FINE-GRAINED ATTRIBUTE ANALYSIS FOR PERSON RE-IDENTIFICATION

Suncheng Xiang, Yuzhuo Fu, Guanjie You, Ting Liu ...................................................................... 3765

MLSP-37: PATTERN RECOGNITION AND CLASSIFICATION 2

UNIFIED CLUSTERING AND OUTLIER DETECTION ON SPECIALIZED HARDWARE

Eldan Cohen, Hayato Ushijima-Mwesigwa, Avradip Mandal, Arnab Roy ...................................... 3770

CLASS-IMBALANCED CLASSIFIERS USING ENSEMBLES OF GAUSSIAN PROCESSES
AND GAUSSIAN PROCESS LATENT VARIABLE MODELS

Liu Yang, Cassandra Heiselman, J. Gerald Quirk, Petar M. Djuric ...................................................... 3775

IMPROVING DEEP LEARNING SOUND EVENTS CLASSIFIERS USING GRAM MATRIX
FEATURE-WISE CORRELATIONS

Antonio Joia Neto, Andre G. C. Pacheco, Diogo Carbonera Luvizon .................................................. 3780

ADVERSARIALY ROBUST CLASSIFICATION BASED ON GLRT

Bhagyashree Puranik, Upamanyu Madhow, Ramtin Pedarsani ........................................................ 3785

CROSS-CORPUS SPEECH EMOTION RECOGNITION USING JOINT DISTRIBUTION
ADAPTIVE REGRESSION

Jiacheng Zhang, Lin Jiang, Yuan Zong, Wenming Zheng, Li Zhao ....................................................... 3790

META-COGNITION-BASED SIMPLE AND EFFECTIVE APPROACH TO OBJECT DETECTION

Sannidhi P Kumar, Chandan Gautam, Suresh Sundaram ................................................................... 3795

MLSP-38: NEURAL NETWORKS FOR CLUSTERING AND CLASSIFICATION

GRAPHNET: GRAPH CLUSTERING WITH DEEP NEURAL NETWORKS

Xianchao Zhang, Jie Mu, Han Liu, Xiaotong Zhang .......................................................................... 3800

ATTENTION ENHANCED SPATIAL TEMPORAL NEURAL NETWORK FOR HRRP RECOGNITION

Yuchen Chu, Zunhua Guo .................................................................................................................. 3805

DHCN: DEEP HIERARCHICAL CONTEXT NETWORKS FOR IMAGE ANNOTATION

Mingyuan Jiu, Hichem Sahbi ............................................................................................................... 3810

ON THE POWER OF DEEP BUT NAIVE PARTIAL LABEL LEARNING

Junghoon Seo, Joon Suk Huh ............................................................................................................. 3820

ADVANCES IN MORPHOLOGICAL NEURAL NETWORKS: TRAINING, PRUNING AND ENFORCING SHAPE CONSTRAINTS

Nikolaos Dimitriadis, Petros Maragos .............................................................................................. 3825
MLSP-39: ADVERSARIAL MACHINE LEARNING

ADVERSARIAL LEARNING VIA PROBABILISTIC PROXIMITY ANALYSIS ................................................................. 3830
Jarrod Hollis, Jinsub Kim, Raviv Raich

CLASS AWARE ROBUST TRAINING .................................................................................................................... 3835
Zhikang Xia, Bin Chen, Tao Dai, Shu-Tao Xia

NON-SINGULAR ADVERSARIAL ROBUSTNESS OF NEURAL NETWORKS ................................................................. 3840
Yu-Lin Tsai, Chia-Yi Hsu, Chia-Mu Yu, Pin-Yu Chen

TOWARDS ADVERSARIAL ROBUSTNESS VIA COMPACT FEATURE REPRESENTATIONS ............................................................ 3845
Muhammad A. Shah, Raphael Olivier, Bhiksha Raj

ADVERSARIAL EXAMPLES DETECTION BEYOND IMAGE SPACE ................................................................. 3850
Kejiang Chen, Yuefeng Chen, Hang Zhou, Chuan Qin, Xiaofeng Mao, Weiming Zhang, Nenghai Yu

STRONG DATA AUGMENTATION SANITIZES POISONING AND BACKDOOR ATTACKS WITHOUT AN ACCURACY TRADEOFF .................................................................................................................. 3855
Eitan Borgnia, Valeriu Cherepanova, Liam Fowl, Amin Ghiasi, Jonas Geiping, Micah Goldblum, Tom Goldstein, Arjun Gupta

MLSP-40: CONTRASTIVE LEARNING

CONTRASTIVE PREDICTIVE CODING SUPPORTED FACTORIZED VARIATIONAL AUTOENCODER FOR UNSUPERVISED LEARNING OF DISENTANGLED SPEECH REPRESENTATIONS .......................................................................................................................... 3860
Janek Ebbers, Michael Kuhlmann, Tobias Cord-Landwehr, Reinhold Haeb-Umbach

CONTRASTIVE SEPARATIVE CODING FOR SELF-SUPERVISED REPRESENTATION LEARNING .......................................................................................................................... 3865
Jun Wang, Max W Y. Lam, Dan Su, Dong Yu

CONTRASTIVE SEMI-SUPERVISED LEARNING FOR ASR ............................................................................................. 3870
Alex Xiao, Christian Fuegen, Abdelrahman Mohamed

CONTRASTIVE LEARNING OF GENERAL-PURPOSE AUDIO REPRESENTATIONS .................. 3875
Aaqib Saeed, David Grangier, Neil Zeghidour

SEQ-CPC: SEQUENTIAL CONTRASTIVE PREDICTIVE CODING FOR AUTOMATIC SPEECH RECOGNITION .......................................................................................................................... 3880
Yulong Chen, Jianping Zhao, Weiqi Wang, Ming Fang, Haimei Kang, Lu Wang, Tao Wei, Jun Ma, Shaojun Wang, Jing Xiao

ON SCALING CONTRASTIVE REPRESENTATIONS FOR LOW-RESOURCE SPEECH RECOGNITION .......................................................................................................................... 3885
Lasse Borgholt, Tycho M. S. Tax, Jakob D. Havnorn, Lars Maaloe, Christian Igel
**MLSP-41: DEEP LEARNING OPTIMIZATION**

CONVEX NEURAL AUTOREGRESSIVE MODELS: TOWARDS TRACTABLE, EXPRESSIVE, AND THEORETICALLY-BACKED MODELS FOR SEQUENTIAL FORECASTING AND GENERATION ................................................................. 3890  
Vikul Gupta, Burak Bartan, Tolga Ergen, Mert Pilanci

INERTIAL PROXIMAL DEEP LEARNING ALTERNATING MINIMIZATION FOR EFFICIENT NEURAL NETWORK TRAINING ....................................................................................................................... 3895  
Linbo Qiao, Tao Sun, Hengyue Pan, Dongsheng Li

KALMAN OPTIMIZER FOR CONSISTENT GRADIENT DESCENT ................................................................. 3900  
Xingyi Yang

KALMANNET: DATA-DRIVEN KALMAN FILTERING ................................................................. 3905  
Guy Revach, Nir Shlezinger, Ruud J. G. Van Sloun, Yonina C. Eldar

HCGM-NET: A DEEP UNFOLDING NETWORK FOR FINANCIAL INDEX TRACKING ................. 3910  
Ruben Pauwels, Evaggelia Tsiligiani, Nikos Deligiannis

AUGMENTING TRANSFERRED REPRESENTATIONS FOR STOCK CLASSIFICATION ............... 3915  
Elizabeth Fons, Paula Dawson, Xian-Jun Zeng, John Keane, Alexandros Iosifidis

**MLSP-42: NEURAL NETWORK PRUNING**

A FRAMEWORK FOR PRUNING DEEP NEURAL NETWORKS USING ENERGY-BASED MODELS ................................................................................................................................. 3920  
Hojjat Salehinejad, Shahrokh Valaee

PROTOTYPE-BASED PERSONALIZED PRUNING ............................................................................. 3925  
Jangho Kim, Simyung Chang, Sungrack Yun, Nojun Kwak

TENSOR REORDERING FOR CNN COMPRESSION ............................................................................. 3930  
Matej Ulicny, Vladimir A. Krylov, Rozenn Dahyot

PRUNING OF CONVOLUTIONAL NEURAL NETWORKS USING ISING ENERGY MODEL .......... 3935  
Hojjat Salehinejad, Shahrokh Valaee

REWEIGHTED DYNAMIC GROUP CONVOLUTION ............................................................................. 3940  
Weiwei Chen, Chong Wang, Zhehao Zhang, Zheng Huo, Linlin Gao

LAYER-WISE INTERPRETATION OF DEEP NEURAL NETWORKS USING IDENTITY INITIALIZATION ................................................................................................................................. 3945  
Shohei Kubota, Hideaki Hayashi, Tomohiro Hayase, Seiichi Uchida

**MLSP-43: BIOMEDICAL APPLICATIONS**

DETECTION OF POST-TRAUMATIC STRESS DISORDER USING LEARNED TIME-FREQUENCY REPRESENTATIONS FROM PUPILLOMETRY ................................................................. 3950  
Bilal Taha, Megan Kirk, Paul Ritvo, Dimitrios Hatzinakos

SUBJECT-IN Variant EEG REPRESENTATION LEARNING FOR EMOTION RECOGNITION ....... 3955  
Soheil Rayatdoost, Yufeng Yin, David Rudrauf, Mohammad Soleymani
TOWARDS PARKINSON’S DISEASE PROGNOSIS USING SELF-SUPERVISED LEARNING AND ANOMALY DETECTION................................................................................................................... 3960
Hongchao Jiang, Wei Yang Bryan Lim, Jer Shyuan Ng, Yu Wang, Ying Chi, Chunyan Miao

IN-BED PRESSURE-BASED POSE ESTIMATION USING IMAGE SPACE REPRESENTATION LEARNING................................................................................................................ 3965
Vandad Davoodnia, Saeed Ghorbani, Ali Etemad

TOWARDS THE DEVELOPMENT OF SUBJECT-INDEPENDENT INVERSE METABOLIC MODELS ................................................................................................................ 3970
Seyedhooman Sajjadi, Anurag Das, Ricardo Gutierrez-Osuna, Theodora Chaspari, Projna Paromita, Laura E. Ruebush, Nicolaas E. Deutz, Bobak J. Mortazavi

HUMAN-EXPERT-LEVEL BRAIN TUMOR DETECTION USING DEEP LEARNING WITH DATA DISTILLATION AND AUGMENTATION.................................................................. 3975
Diyuan Lu, Nenad Polomac, Iskra Gacheva, Elke Hattingen, Jochen Triesch

MLSP-44: MULTIMODAL DATA AND APPLICATIONS

MULTIMODAL PUNCTUATION PREDICTION WITH CONTEXTUAL DROPOUT .............................................................................................................................................. 3980
Andrew Silva, Barry-John Theobald, Nicholas Apostoloff

MULTI-MODAL LABEL DEQUANTIZED GAUSSIAN PROCESS LATENT VARIABLE MODEL FOR ORDINAL LABEL ESTIMATION ............................................................................ 3985
Masanao Matsumoto, Keisuke Maeda, Naoki Saito, Takahiro Ogawa, Miki Haseyama

GENERATIVE INFORMATION FUSION............................................................................................................................................................................................... 3990
Kenneth Tran, Wesam Sakla, Hamid Krim

SELF-AUGMENTED MULTI-MODAL FEATURE EMBEDDING ......................................................................................................................................................... 3995
Shinnosuke Matsuo, Seiichi Uchida, Brian Kenji Iwana

OPTIMIZE WHAT MATTERS: TRAINING DNN-HMM KEYWORD SPOTTING MODEL USING END METRIC........................................................................................................... 4000
Ashish Shrivastava, Arnav Kundu, Chandra Dhir, Devang Naik, Oncel Tuzel

CO-ATTENTIONAL TRANSFORMERS FOR STORY-BASED VIDEO UNDERSTANDING ................................................. 4005
Björn Bebensee, Byoung-Tak Zhang

MLSP-45: PERFORMANCE BOUNDS

DEEP GENERATIVE DEMIXING: ERROR BOUNDS FOR DEMIXING SUBGAUSSIAN MIXTURES OF LIPSCHITZ SIGNALS ........................................................................................................................................ 4010
Aaron Berk

TOWARDS AN INTRINSIC DEFINITION OF ROBUSTNESS FOR A CLASSIFIER ...................................................................................................................... 4015
Théo Giraudon, Vincent Gripon, Matthias Löwe, Franck Vermet

PHASE TRANSITIONS FOR ONE-VS-ONE AND ONE-VS-ALL LINEAR SEPARABILITY IN MULTICLASS GAUSSIAN MIXTURES ................................................................................................................................. 4020
Ganesh Ramachandra Kini, Christos Thrampoulidis

LEAKY INTEGRATOR DYNAMICAL SYSTEMS AND REACHABLE SETS ......................................................................................................................... 4025
Brian Whiteaker, Peter Gerstoft
<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>BENIGN OVERFITTING IN BINARY CLASSIFICATION OF GAUSSIAN MIXTURES</td>
<td>4030</td>
</tr>
<tr>
<td>Ke Wang, Christos Thrampoulidis</td>
<td></td>
</tr>
<tr>
<td>AN ORDER-OPTIMAL ADAPTIVE TEST PLAN FOR NOISY GROUP TESTING UNDER</td>
<td>4035</td>
</tr>
<tr>
<td>UNKNOWN NOISE MODELS</td>
<td></td>
</tr>
<tr>
<td>Sudeep Salgia, Qing Zhao</td>
<td></td>
</tr>
<tr>
<td>MLSP-46: THEORY AND APPLICATIONS</td>
<td></td>
</tr>
<tr>
<td>SAPAUGMENT: LEARNING A SAMPLE ADAPTIVE POLICY FOR DATA AUGMENTATION</td>
<td>4040</td>
</tr>
<tr>
<td>Ting-Yao Hu, Ashish Shrivastava, Jen-Hao Rick Chang, Hema Koppula,</td>
<td></td>
</tr>
<tr>
<td>Stefan Braun, Kyuyeon Hwang, Ozlem Kalinli, Oncel Tuzel</td>
<td></td>
</tr>
<tr>
<td>HIERARCHICAL CODED ELASTIC COMPUTING</td>
<td>4045</td>
</tr>
<tr>
<td>Shahrzad Kiani, Tharindu Adikari, Stark C. Draper</td>
<td></td>
</tr>
<tr>
<td>SYNTHESIZE &amp; LEARN: JOINTLY OPTIMIZING GENERATIVE AND CLASSIFIER</td>
<td>4050</td>
</tr>
<tr>
<td>NETWORKS FOR IMPROVED DROWSINESS DETECTION</td>
<td></td>
</tr>
<tr>
<td>Sandipan Banerjee, Aijen Joshi, Ahmed Ghoneim, Surv Kyal, Taniya</td>
<td></td>
</tr>
<tr>
<td>Mishra</td>
<td></td>
</tr>
<tr>
<td>A JOINT CONVOLUTIONAL AND SPATIAL QUAD-DIRECTIONAL LSTM NETWORK FOR</td>
<td>4055</td>
</tr>
<tr>
<td>PHASE UNWRAPPING</td>
<td></td>
</tr>
<tr>
<td>Malsha V. Perera, Ashwin De Silva</td>
<td></td>
</tr>
<tr>
<td>MLSP-47: APPLICATIONS OF MACHINE LEARNING</td>
<td></td>
</tr>
<tr>
<td>INTEGRATED CLASSIFICATION AND LOCALIZATION OF TARGETS USING BAYESIAN</td>
<td>4060</td>
</tr>
<tr>
<td>FRAMEWORK IN AUTOMOTIVE RADARS</td>
<td></td>
</tr>
<tr>
<td>Anand Dubey, Avik Santra, Jonas Fuchs, Maximilian Lübke, Robert</td>
<td></td>
</tr>
<tr>
<td>Weigel, Fabian Lurz</td>
<td></td>
</tr>
<tr>
<td>A DNN AUTOENCODER FOR AUTOMOTIVE RADAR INTERFERENCE MITIGATION</td>
<td>4065</td>
</tr>
<tr>
<td>Shengyi Chen, Jalal Taghia, Tai Fei, Uwe Kühnau, Nils Pohl, Rainer</td>
<td></td>
</tr>
<tr>
<td>Martin</td>
<td></td>
</tr>
<tr>
<td>DURAS: DEEP UNFOLDED RADAR SENSING USING DOPPLER FOCUSING</td>
<td>4070</td>
</tr>
<tr>
<td>Pranav Goyal, Satish Mulleti, Anubha Gupta, Yonina C. Eldar</td>
<td></td>
</tr>
<tr>
<td>NNAKF: A NEURAL NETWORK ADAPTED KALMAN FILTER FOR TARGET TRACKING</td>
<td>4075</td>
</tr>
<tr>
<td>Sami Jouaber, Silvère Bonnabel, Santiago Velasco-Forero, Marion</td>
<td></td>
</tr>
<tr>
<td>Piltié</td>
<td></td>
</tr>
<tr>
<td>MULTI-SAMPLE ONLINE LEARNING FOR SPIKING NEURAL NETWORKS BASED ON</td>
<td>4080</td>
</tr>
<tr>
<td>GENERALIZED EXPECTATION MAXIMIZATION</td>
<td></td>
</tr>
<tr>
<td>Hyeryung Jang, Osvaldo Simeone</td>
<td></td>
</tr>
<tr>
<td>PROBABILISTIC GRAPH NEURAL NETWORKS FOR TRAFFIC SIGNAL CONTROL</td>
<td>4085</td>
</tr>
<tr>
<td>Ting Zhong, Zheyang Xu, Fan Zhou</td>
<td></td>
</tr>
<tr>
<td>MLSP-48: NEURAL NETWORK APPLICATIONS</td>
<td></td>
</tr>
<tr>
<td>TASK-AWARE NEURAL ARCHITECTURE SEARCH</td>
<td>4090</td>
</tr>
<tr>
<td>Cat P. Le, Mohammadreza Soltani, Robert Ravier, Vahid Taroikh</td>
<td></td>
</tr>
</tbody>
</table>
F-NET: FUSION NEURAL NETWORK FOR VEHICLE TRAJECTORY PREDICTION IN AUTONOMOUS DRIVING ................................................................. 4095
Jue Wang, Ping Wang, Chao Zhang, Kuifeng Su, Jun Li

UNSUPERVISED RECONSTRUCTION OF SEA SURFACE CURRENTS FROM AIS MARITIME TRAFFIC DATA USING LEARNABLE VARIATIONAL MODELS ........................................... 4100
S. Benaïchouche, C. Le Goff, Y. Guichoux, F. Rousseau, R. Fablet

A COMPACT JOINT DISTILLATION NETWORK FOR VISUAL FOOD RECOGNITION .................... 4105
Zhao Heng, Kim-Hui Yap, Alex Chichung Kot

PIPELINE SAFETY EARLY WARNING METHOD FOR DISTRIBUTED SIGNAL USING BILINEAR CNN AND LIGHTGBM ................................................................. 4110
Yiyuan Yang, Yi Li, Haifeng Zhang

DEEP LEARNING BASED HYBRID PRECODING IN DUAL-BAND COMMUNICATION SYSTEMS ................................................................................. 4115
Rafail Ismayilov, Renato L. G. Cavalcante, Slawomir Stanczak

DEEP LEARNING-BASED CROSS-LAYER RESOURCE ALLOCATION FOR WIRED COMMUNICATION SYSTEMS ......................................................... 4120
Pourya Behmandpoor, Jeroen Verdyck, Marc Moonen

MMSP-1: MULTIMEDIA SIGNAL PROCESSING

ATVIO: ATTENTION GUIDED VISUAL-INERTIAL ODOMETRY ......................................................... 4125
Li Liu, Ge Li, Thomas H Li

FEATURE INTEGRATION VIA SEMI-SUPERVISED ORDINALLY MULTI-MODAL GAUSSIAN PROCESS LATENT VARIABLE MODEL .............................................. 4130
Kyohei Kamikawa, Keisuke Maeda, Takahiro Ogawa, Miki Haseyama

A MULTI-LAYER MULTI-CHANNEL ATTENTIVE NETWORK FOR GENDER AND AGE RECOGNITION ................................................................. 4135
Jia Chen, Haiping Yu, Yimei Kang

EFFECT OF LANGUAGE PROFICIENCY ON SUBJECTIVE EVALUATION OF NOISE SUPPRESSION ALGORITHMS ................................................................. 4140
Babak Naderi, Gabriel Mittag, Rafael Zequeira Jiménez, Sebastian Möller

MELODY HARMONIZATION USING ORDERLESS NADE, CHORD BALANCING, AND BLOCKED GIBBS SAMPLING ................................................. 4145
Chung-En Sun, Yi-Wei Chen, Hung-Shin Lee, Yen-Hsing Chen, Hsin-Min Wang

MMSP-2: DEEP LEARNING FOR MULTIMEDIA ANALYSIS AND PROCESSING

CROSS-DOMAIN SEMI-SUPERVISED DEEP METRIC LEARNING FOR IMAGE SENTIMENT ANALYSIS ............................................................................. 4150
Yun Liang, Keisuke Maeda, Takahiro Ogawa, Miki Haseyama

AUDIOVISUAL HIGHLIGHT DETECTION IN VIDEOS ..................................................................... 4155
Karel Mundnich, Alexandra Fenster, Aparna Khare, Shiva Sundaram
TEACHER-ASSISTED MINI-BATCH SAMPLING FOR BLIND DISTILLATION USING METRIC LEARNING .................................................. 4160
Nakamasa Inoue

RULE-EMBEDDED NETWORK FOR AUDIO-VISUAL VOICE ACTIVITY DETECTION IN LIVE MUSICAL VIDEO STREAMS .................................................................................................................................... 4165
Yuanbo Hou, Yi Deng, Bilei Zhu, Zejun Ma, Dick Botteldooren

REINFORCEMENT STACKED LEARNING WITH SEMANTIC-ASSOCIATED ATTENTION FOR VISUAL QUESTION ANSWERING ........................................................................................................... 4170
Xinyu Xiao, Chunxia Zhang, Shiming Xiang, Chunhong Pan

HIERARCHICAL REFINED ATTENTION FOR SCENE TEXT RECOGNITION ............................................................ 4175
Min Zhang, Meng Ma, Ping Wang

MMSP-3: MULTIMEDIA SYNTHESIS AND ENHANCEMENT

COLLABORATIVE LEARNING TO GENERATE AUDIO-VIDEO JOINTLY ........................................................................ 4180
Vinod K Kurmi, Vipul Bajaj, Badri N Patro, KS Venkatesh, Vinay P Namboodiri, Preethi Jyothi

AN ATTENTION-SEQ2SEQ MODEL BASED ON CRNN ENCODING FOR AUTOMATIC LABANOTATION GENERATION FROM MOTION CAPTURE DATA ........................................................................................................... 4185
Min Li, Zhenjiang Miao, Xiao-Ping Zhang, Wanru Xu

SHOW AND SPEAK: DIRECTLY SYNTHESIZE SPOKEN DESCRIPTION OF IMAGES ........................................................................ 4190
Xinsheng Wang, Siyuan Feng, Jihua Zhu, Mark Hasegawa-Johnson, Odette Scharenborg

DRAWGAN: TEXT TO IMAGE SYNTHESIS WITH DRAWING GENERATIVE ADVERSARIAL NETWORKS ..................................................................................................................... 4195
Zhiqiang Zhang, Jinjia Zhou, Wenxin Yu, Ning Jiang

DISENTANGLING SUBJECT-DEPENDENT/-INDEPENDENT REPRESENTATIONS FOR 2D MOTION RETARGETING ........................................................................................................................... 4200
Fanglu Xie, Go Irie, Tatsushi Matsubayashi

MMSP-4: IMAGE, VIDEO AND POINT CLOUD CODING

NETWORK AND CONTENT-DEPENDENT BITRATE LADDER ESTIMATION FOR ADAPTIVE BITRATE VIDEO STREAMING ............................................................................................................................. 4205
Pierre Lebreton, Kazuhiisa Yamagishi

ULTRA-LOW BITRATE VIDEO CONFERENCING USING DEEP IMAGE ANIMATION ........................................................................ 4210
Goluck Konuko, Giuseppe Valenzise, Stéphane Lathuilière

HIERARCHICAL BIT-WISE DIFFERENTIAL CODING (HBDC) OF POINT CLOUD ATTRIBUTES ............................................................................................................................. 4215
Yan Huang, Bin Wang, C.-C. Jay Kuo, Hui Yuan, Jingliang Peng

LEARNING-BASED LOSSLESS COMPRESSION OF 3D POINT CLOUD GEOMETRY ........................................................................ 4220
Dat Thanh Nguyen, Maurice Quach, Giuseppe Valenzise, Pierre Duhamel

IMAGE CODING WITH NEURAL NETWORK-BASED COLORIZATION ........................................................................ 4225
Diogo Lopes, João Ascenso, Catarina Brites, Fernando Pereira
JOINT REINFORCEMENT LEARNING AND GAME THEORY BITRATE CONTROL METHOD FOR 360-DEGREE DYNAMIC ADAPTIVE STREAMING .......................................................... 4230
Xuekai Wei, Mingliang Zhou, Sam Kwong, Hui Yuan, Tao Xiang

MMSP-5: HUMAN CENTRIC MULTIMEDIA 1

HCAG: A HIERARCHICAL CONTEXT-AWARE GRAPH ATTENTION MODEL FOR DEPRESSION DETECTION......................................................................................................................... 4235
Meng Niu, Kai Chen, Qingcai Chen, Lufeng Yang

WHEN FACE RECOGNITION MEETS OCCLUSION: A NEW BENCHMARK ........................................... 4240
Baojin Huang, Zhongyuan Wang, Guangcheng Wang, Kui Jiang, Kangli Zeng, Zhen Han, Xin Tian, Yuhong Yang

A TRIPLET APPEARANCE PARSING NETWORK FOR PERSON RE-IDENTIFICATION .................. 4245
Mingfu Xiong, Zhongyuan Wang, Ruhan He, Xinrong Hu, Ming Cheng, Xiao Qin, Jia Chen

PART-ALIGNED NETWORK WITH BACKGROUND FOR MISALIGNED PERSON SEARCH ........... 4250
Xian Zhong, Yiting Liu, Wenxin Huang, Xiao Wang, Bo Ma, Jingleing Yuan

LEARNING POSE-ADAPTIVE LIP SYNC WITH CASCADED TEMPORAL CONVOLUTIONAL NETWORK ................................................................................................................. 4255
Ruobing Zheng, Bo Song, Changjiang Ji

ASSESSMENT OF BIPOLAR DISORDER USING HETEROGENEOUS DATA OF SMARTPHONE-BASED DIGITAL PHENOTYPING ................................................................................. 4260

MMSP-6: HUMAN CENTRIC MULTIMEDIA 2

MULTI-GRANULARITY FEATURE INTERACTION AND RELATION REASONING FOR 3D DENSE ALIGNMENT AND FACE RECONSTRUCTION ........................................... 4265
Lei Li, Xiangzheng Li, Kangbo Wu, Kui Lin, Suping Wu

INDEPENDENT SIGN LANGUAGE RECOGNITION WITH 3D BODY, HANDS, AND FACE RECONSTRUCTION ............................................................................................................... 4270
Agelos Kratimenos, Georgios Pavlakos, Petros Maragos

MULTIMODAL CROSS- AND SELF-ATTENTION NETWORK FOR SPEECH EMOTION RECOGNITION ..................................................................................................................... 4275
Licai Sun, Bin Liu, Jianhua Tao, Zheng Lian

MULTI-TARGET DOA ESTIMATION WITH AN AUDIO-VISUAL FUSION MECHANISM .............. 4280
Xinyuan Qian, Maulik Madhavi, Zexu Pan, Jiadong Wang, Haizhou Li

IMPROVING MULTIMODAL SPEECH ENHANCEMENT BY INCORPORATING SELF-SUPERVISED AND CURRICULUM LEARNING ............................................................................. 4285
Ying Cheng, Mengyu He, Jiashuo Yu, Rui Feng

AUTOENCODER FOR VIBROTACTILE SIGNAL COMPRESSION .................................................... 4290
Zhuoran Li, Rania Hassen, Zhou Wang
MMSP-7: MULTIMODAL PERCEPTION, INTEGRATION AND MULTISENSORY FUSION

CROSS-MODAL KNOWLEDGE DISTILLATION FOR FINE-GRAINED ONE-SHOT CLASSIFICATION ................................................................. 4295
Jiabao Zhao, Xin Lin, Yifan Yang, Jing Yang, Liang He

LEARNING AUDIO-VISUAL CORRELATIONS FROM VARIATIONAL CROSS-MODAL GENERATION .......................................................... 4300
Ye Zhu, Yu Wu, Hugo Latapie, Yi Yang, Yan Yan

ECCL: EXPLICIT CORRELATION-BASED CONVOLUTIONAL BOUNDARY LOCATOR FOR MOMENT LOCALIZATION .................................................. 4305
Xinfang Liu, Xiushan Nie, Junya Teng, Fanchang Hao, Yilong Yin

COOPNET: MULTI-MODAL COOPERATIVE GENDER PREDICTION IN SOCIAL MEDIA USER PROFILING .................................................... 4310
Lin Li, Kaixi Hu, Yunpei Zheng, Jianquan Liu, Kong Aik Lee

ROBUST LATENT REPRESENTATIONS VIA CROSS-MODAL TRANSLATION AND ALIGNMENT .................................................................. 4315
Vandana Rajan, Alessio Brutti, Andrea Cavallaro

SEMI-SUPERVISED MULTIMODAL IMAGE TRANSLATION FOR MISSING MODALITY IMPUTATION .......................................................... 4320
Wangbin Sun, Fei Ma, Yang Li, Shao-Lun Huang, Shiguang Ni, Lin Zhang

MMSP-8: MULTIMEDIA RETRIEVAL AND SIGNAL DETECTION

DEEP ADVERSARIAL QUANTIZATION NETWORK FOR CROSS-MODAL RETRIEVAL ............................................................. 4325
Yu Zhou, Yong Feng, Mingliang Zhou, Baohua Qiang, Leong Hou U, Jiajie Zhu

SCALABLE DISCRIMINATIVE DISCRETE HASHING FOR LARGE-SCALE CROSS-MODAL RETRIEVAL .................................................. 4330
Jianyang Qin, Lunke Fei, Jian Zhu, Jie Wen, Chunwei Tian, Shuai Wu

HIERARCHICAL SIMILARITY LEARNING FOR LANGUAGE-BASED PRODUCT IMAGE RETRIEVAL .......................................................... 4335
Zhe Ma, Fenghao Liu, Jianfeng Dong, Xiaoye Qu, Yuan He, Shouling Ji

BIDIRECTIONAL FOCUSED SEMANTIC ALIGNMENT ATTENTION NETWORK FOR CROSS-MODAL RETRIEVAL ........................................ 4340
Shuli Cheng, Liejun Wang, Anyu Du, Yongming Li

DETECTION OF AUDIO-VIDEO SYNCHRONIZATION ERRORS VIA EVENT DETECTION .............................................................. 4345
Joshua P. Ebeneze, Yongjun Wu, Hai Wei, Sriram Sethuraman, Zongyi Liu

FC^2RN: A FULLY CONVOLUTIONAL CORNER REFINEMENT NETWORK FOR ACCURATE MULTI-ORIENTED SCENE TEXT DETECTION ............................................................. 4350
Xugong Qin, Yu Zhou, Youhui Guo, Dayan Wu, Weiping Wang
SAM-1: DIRECTION OF ARRIVAL ESTIMATION 1

DOA ESTIMATION OF A HIDDEN RF SOURCE EXPLOITING SIMPLE BACKSCATTER RADIO TAGS................................................................. 4355
Georgios Vougioukas, Aggelos Bletsas

PROBABILITY OF RESOLUTION OF G-MUSIC: AN ASYMPTOTIC APPROACH .................................................. 4360
David Schenck, Xavier Mestre, Marius Pesavento

A PARTIALLY-RELAXED ROBUST DOA ESTIMATOR UNDER NON-GAUSSIAN LOW-RANK INTERFERENCE AND NOISE .................................................. 4365
Minh Trinh-Hoang, Mohammed Nabil El Korso, Marius Pesavento

NON-COHERENT DOA ESTIMATION OF OFF-GRID SIGNALS WITH UNIFORM CIRCULAR ARRAYS................................................................. 4370
Zhengyu Wan, Wei Liu

ENHANCED STANDARD ESPRIT FOR OVERCOMING IMPERFECTIONS IN DOA ESTIMATION................................................................. 4375
Majdoddin Esfandiari, Sergiy A. Vorobyov

SAM-2: DIRECTION OF ARRIVAL ESTIMATION 2

CONSTRAINED TENSOR DECOMPOSITION FOR 2D DOA ESTIMATION IN TRANSMIT BEAMSPACE MIMO RADAR WITH SUBARRAYS.................................................. 4380
Feng Xu, Sergiy A. Vorobyov

ALTERNATING PROJECTIONS GRIDLESS COVARIANCE-BASED ESTIMATION FOR DOA ................................................................. 4385
Yongsung Park, Peter Gerstoft

SYNTHETIC DATA FOR DNN-BASED DOA ESTIMATION OF INDOOR SPEECH .................................................. 4390
Femke B. Gelderblom, Yi Liu, Johannes Kvam, Tor Andre Myrvoll

DIRECTION OF ARRIVAL ESTIMATION FOR NON-COHERENT SUB-ARRAYS VIA JOINT SPARSE AND LOW-RANK SIGNAL RECOVERY .................................................. 4395
Tom Tirer, Oded Bialer

SPARSITY AND NONNEGATIVITY CONSTRAINED KRYLOV APPROACH FOR DIRECTION OF ARRIVAL ESTIMATION .................................................. 4400
Hamza Baali, Abdesselam Bouzerdoum, Abdelkrim Khelif

SAM-3: MIMO RADAR ARRAY PROCESSING

HYBRID ANALOG-DIGITAL MIMO RADAR RECEIVERS WITH BIT-LIMITED ADCS.................................................. 4405
Feng Xi, Nir Shlezinger, Yonina C. Eldar

SPARSE ARRAY TRANSCEIVER DESIGN FOR ENHANCED ADAPTIVE BEAMFORMING IN MIMO RADAR .................................................. 4410
Syed A. Hamza, Weitong Zhai, Xiangrong Wang, Moeness G. Amin

SPARSE PARAMETER ESTIMATION FOR PMCW MIMO RADAR USING FEW-BIT ADCS .................................................. 4415
Chao-Yi Wu, Jian Li, Tan F. Wong
PARAMETER IDENTIFIABILITY OF SPATIAL-SMOOTHING-BASED BISTATIC MIMO RADAR .............................................................. 4420
Junpeng Shi, Fangqing Wen, Yongxiang Liu, Qimmu Shen, Zhihui Li, Zhen Liu

PARAMETER ESTIMATION FOR COHERENT PASSIVE MIMO RADAR WITH UNKNOWN SIGNALS UNDER DIRECT PATH INFLUENCE ........................................................................ 4425
Zhen Wang, Qian He

RIEMANNIAN GEOMETRIC OPTIMIZATION METHODS FOR JOINT DESIGN OF TRANSMIT SEQUENCE AND RECEIVE FILTER OF MIMO RADAR .......................................................... 4430
Jie Li, Guisheng Liao, Yan Huang, Arye Nehorai

SAM-4: MIMO AND MASSIVE MIMO ARRAY PROCESSING

HIGH ACCURACY TRACKING OF TARGETS USING MASSIVE MIMO .................................................................................. 4435
Xiaolu Zeng, Feng Zhang, Beibei Wang, K. J. Ray Liu

ADMM-BASED FAST ALGORITHM FOR ROBUST MULTI-GROUP MULTICAST BEAMFORMING ................................................................................................................................. 4440
Niloofar Mohamadi, Min Dong, Shahram Shahbazpanahi

SCALABLE AND DISTRIBUTED MMSE ALGORITHMS FOR UPLINK RECEIVE COMBINING IN CELL-FREE MASSIVE MIMO SYSTEMS ........................................................................... 4445
Robbe Van Rompaey, Marc Moonen

ANTENNA SELECTION FOR MASSIVE MIMO SYSTEMS BASED ON POMDP FRAMEWORK ................................................................................................................................. 4450
Sara Sharifi, Shahram Shahbaz Panahi, Min Dong

RIS-AIDED JOINT LOCALIZATION AND SYNCHRONIZATION WITH A SINGLE-ANTENNA MMWAVE RECEIVER ................................................................................................................... 4455
Alessio Fascista, Angelo Coluccia, Henk Wymeersch, Gonzalo Seco-Granados

JOINT CHANNEL, DATA, AND PHASE-NOISE ESTIMATION IN MIMO-OFDM SYSTEMS USING A TENSOR MODELING APPROACH .............................................................................. 4460
Bruno Sokal, Paulo R. B. Gomes, André L. F. De Almeida, Martin Haardt

SAM-5: MICROPHONE ARRAY SIGNAL PROCESSING

ROBUST STEERABLE DIFFERENTIAL BEAMFORMERS WITH NULL CONSTRAINTS FOR CONCENTRIC CIRCULAR MICROPHONE ARRAYS .................................................................................. 4465
Xuehan Wang, Gongping Huang, Israel Cohen, Jacob Benesty, Jingdong Chen

CLOSE-TALKING RECORDING WITH PLANARLY DISTRIBUTED MICROPHONES .......................................................................... 4470
Takuma Okamoto

(W)EARABLE MICROPHONE ARRAY AND ULTRASONIC ECHO LOCALIZATION FOR COARSE INDOOR ENVIRONMENT MAPPING .......................................................................................... 4475
Felix Pfrendtner, Jing Yang, Gábor Sörös

CHARACTERIZATION OF MEMS MICROPHONE SENSITIVITY AND PHASE DISTRIBUTIONS WITH APPLICATIONS IN ARRAY PROCESSING ........................................................................... 4480
Patrick W. A. Wijnings, Sander Stuijk, Rick Scholte, Henk Corporaal
DIRECTIONAL SPARSE FILTERING USING WEIGHTED LEHMER MEAN FOR BLIND
SEPARATION OF UNBALANCED SPEECH MIXTURES .......................................................... 4485
Karn Watcharasupat, Anh H. T. Nguyen, Ching-Hui Ooi, Andy W. H. Khong

DISTRIBUTED SPEECH SEPARATION IN SPATIALLY UNCONSTRAINED MICROPHONE
ARRAYS ........................................................................................................................................ 4490
Nicolas Furnon, Romain Serizel, Irina Illina, Slim Essid

SAM-6: BEAMFORMING

AN ADAPTIVE NON-LINEAR PROCESS FOR UNDER-DETERMINED VIRTUAL
MICROPHONE BEAMFORMING .............................................................................................. 4495
Mehdi Bekrani, Anh H. T. Nguyen, Andy W. H. Khong

WINDOW BEAMFORMER FOR SPARSE CONCENTRIC CIRCULAR ARRAY .................... 4500
Rajib Sharma, Israel Cohen, Baruch Berdugo

SINGLE-POINT ARRAY RESPONSE CONTROL WITH MINIMUM PATTERN DEVIATION .... 4505
Xiaoyu Ai, Lu Gan

FOCUSING-BASED WIDEBAND ADAPTIVE BEAMFORMING USING COVARIANCE
MATRIX RECONSTRUCTION ................................................................................................. 4510
Peng Chen, Wei Wang, Jingjie Gao

SAM-7: DETECTION AND ESTIMATION 1

BAYESIAN MULTIPLE CHANGE-POINT DETECTION OF PROPAGATING EVENTS .......... 4515
Topi Halme, Eyal Nitzan, Visa Koivunen

ONE-BIT AUTOCORRELATION ESTIMATION WITH NON-ZERO THRESHOLDS .......... 4520
Chun-Lin Liu, Zi-Min Lin

A NOVEL BAYESIAN APPROACH FOR THE TWO-DIMENSIONAL HARMONIC
RETRIEVAL PROBLEM ......................................................................................................... 4525
Rohan R. Pote, Bhaskar D. Rao

ON OVERFITTING IN DISCRETE SUPER-RESOLUTION RECOVERY ......................... 4530
Wenzhe Lu, Heng Qiao

SIML: SIEVED MAXIMUM LIKELIHOOD FOR ARRAY SIGNAL PROCESSING .......... 4535
Matthieu Simeoni, Paul Hurley

ESTIMATION OF GROUNDWATER STORAGE VARIATIONS IN INDUS RIVER BASIN
USING GRACE DATA ........................................................................................................... 4540
Yahya Sattar, Zubair Khalid

SAM-8: DETECTION AND ESTIMATION 2

TEMPORAL EXEMPLAR CHANNELS IN HIGH-MULTIPATH ENVIRONMENTS ............ 4545
Mohamed Kashef, Peter Vouras, Robert Jones, Richard Candell, Kate A. Remley

MULTI-VEHICLE VELOCITY ESTIMATION USING IEEE 802.11AD WAVEFORM .......... 4550
Geonho Han, Sucheol Kim, Junil Choi
REAL-TIME INTERAURAL TIME DELAY ESTIMATION VIA ONSET DETECTION ........................................... 4555
Elizabeth Ren, Gustavo Cid Ornelas, Hans-Andrea Loeliger

EKFNET: LEARNING SYSTEM NOISE STATISTICS FROM MEASUREMENT DATA .......................... 4560
Liang Xu, Ruixin Niu

SLIDING-CAPON BASED CONVOLUTIONAL BEAMSPACE FOR LINEAR ARRAYS ....................... 4565
Po-Chih Chen, P P Vaidyanathan

SAM-9: DETECTION AND CLASSIFICATION
TARGET DETECTION FROM DISTRIBUTED PASSIVE SENSORS: SEMI-LABELED DATA QUANTIZATION .......................................................................................................................................... 4570
Zachariah Sutton, Peter Willett, Stefano Marano

SPARSE FACTORIZATION-BASED DETECTION OF OFF-THE-GRID MOVING TARGETS USING FM CW RADARS .................................................................................................................. 4575
Gilles Monnoyer De Galland, Thomas Feuillen, Luc Vandendorpe, Laurent Jacques

A ROBUST COPULA MODEL FOR RADAR-BASED LANDMINE DETECTION ....................................... 4580
Afief D. Pambudi, Fauzia Ahmad, Abdelhak M. Zoubir

RADAR CLUTTER CLASSIFICATION USING EXPECTATION-MAXIMIZATION METHOD ............... 4585
Sudan Han, Pia Addabbo, Danilo Orlando, Giuseppe Ricci

A META-LEARNING FRAMEWORK FOR FEW-SHOT CLASSIFICATION OF REMOTE SENSING SCENE ......................................................................................................................................... 4590
Pei Zhang, Yunpeng Bai, Dong Wang, Bendu Bai, Ying Li

DIFFERENTIAL CONVOLUTION FEATURE GUIDED DEEP MULTI-SCALE MULTIPLE INSTANCE LEARNING FOR AERIAL SCENE CLASSIFICATION ........................................................ 4595
Beichen Zhou, Jingjun Yi, Qi Bi

SAM-10: SPARSE ARRAY DESIGN AND PROCESSING
GENERALIZED THINNED COPRIME ARRAY FOR DOA ESTIMATION ............................................... 4600
Junpeng Shi, Yongxiang Liu, Fangqing Wen, Zhen Liu, Panhe Hu, Zhenghui Gong

TCLA ARRAY: A NEW SPARSE ARRAY DESIGN WITH LESS MUTUAL COUPLING .......................... 4605
Ahmed M. A. Shaalan, Jun Du, Yan-Hui Tu

LOW MUTUAL COUPLING SPARSE ARRAY DESIGN USING ULA FITTING ........................................ 4610
Wanlu Shi, Yingsong Li, Sergiy A. Vorobyov

LOW-RANK AND SPARSE DECOMPOSITION FOR JOINT DOA ESTIMATION AND CONTAMINATED SENSORS DETECTION WITH SPARSELY CONTAMINATED ARRAYS ........................................ 4615
Huiping Huang, Abdelhak M. Zoubir

FUNDAMENTAL TRADE-OFFS IN NOISY SUPER-RESOLUTION WITH SYNTHETIC APERTURES .......................................................................................................................................... 4620
Sina Shahsavari, Jacob Millhiser, Piya Pal
**SAM-11: ARRAY CALIBRATION AND PERFORMANCE ANALYSIS**

ENHANCED BLIND CALIBRATION OF UNIFORM LINEAR ARRAYS WITH ONE-BIT QUANTIZATION BY KULLBACK-LEIBLER DIVERGENCE COVARIANCE FITTING ................................................. 4625  
_Amir Weiss, Arie Yeredor_

NON-ITERATIVE BLIND CALIBRATION OF NESTED ARRAYS WITH ASYMPTOTICALLY OPTIMAL WEIGHTING ........................................................................................................... 4630  
_Amir Weiss, Arie Yeredor_

SENSOR NETWORKS TDOA SELF-CALIBRATION: 2D COMPLEXITY ANALYSIS AND SOLUTIONS ................................................................................................................................. 4635  
_Luca Ferranti, Kalle Åström, Magnus Oskarsson, Jani Boutellier, Juho Kannala_

FAST AND ROBUST STRATIFIED SELF-CALIBRATION USING TIME-DIFFERENCE-OF-ARRIVAL MEASUREMENTS ................................................................. 4640  
_Martin Larsson, Gabrielle Flood, Magnus Oskarsson, Kalle Åström_

STABILITY ANALYSIS OF THE RC-PLMS ADAPTIVE BEAMFORMER USING A SIMPLE TRANSFER FUNCTION APPROXIMATION ......................................................... 4645  
_Ghattas Akkad, Ali Mansour, Bachar El Hassan, Elie Inaty_

ON THE ASYMPTOTIC PERFORMANCE OF ONE-BIT CO-ARRAY-BASED MUSIC ......................................................... 4650  
_Saeid Sedighi, Bhavani Shankar, Mojtaba Soltanalian, Björn Ottersten_

**SAM-12: TRACKING AND LOCALIZATION**

KLD MINIMIZATION-BASED CONSTRAINED MEASUREMENT FILTERING FOR TWO-STEP TDOA INDOOR TRACKING ................................................................................. 4655  
_Rui Huang, Le Yang, Jun Tao, Yanbo Xue_

A CORRENTROPY BASED ALGORITHM FOR ROBUST LOCALIZATION IN WIRELESS NETWORKS ....................................................................................................................... 4660  
_Mahboobeh Sedighizad, Babak Seyfe, Shahrokh Valaee_

MUG: A MULTIPATH-EXPLOITED AND GRID-FREE LOCALISATION METHOD ............................................................................................... 4665  
_Hengyan Liu, Wei Dai, Yuan Shen_

SPARSE BAYESIAN LEARNING FOR ACOUSTIC SOURCE LOCALIZATION ................................................................................................................................. 4670  
_Ruchi Pandey, Santosh Nannuru, Aditya Siripuram_

AUTOMATIC FINE-GRAINED LOCALIZATION OF UTILITY POLE LANDMARKS ON DISTRIBUTED ACOUSTIC SENSING TRACES BASED ON BILINEAR RESNETS ......................................................... 4675  
_You Lu, Yue Tian, Shaobo Han, Eric Cosatto, Sarper Ozharar, Yangmin Ding_

SSLIDE: SOUND SOURCE LOCALIZATION FOR INDOORS BASED ON DEEP LEARNING ......................................................... 4680  
_Yifan Wu, Roshan Ayyalasomayajula, Michael J. Bianco, Dinesh Bharadia, Peter Gerstoft_

**SAM-13: MULTI-CHANNEL DATA FUSION AND PROCESSING**

PHYSICAL-LAYER SECURITY VIA DISTRIBUTED BEAMFORMING IN THE PRESENCE OF ADVERSARIES WITH UNKNOWN LOCATIONS ......................................................... 4685  
_Yagiz Savas, Abolfazl Hashemi, Abraham P. Vinod, Brian M. Sadler, Ufuk Topcu_
CANONICAL POLYADIC TENSOR DECOMPOSITION WITH LOW-RANK FACTOR MATRICES .................................................................................................................................................... 4690
Anh-Huy Phan, Petr Tichavský, Konstantin Sobolev, Konstantin Sozykin, Dmitry Ermilov, Andrzej Cichocki

A DIFFUSION FXLMS ALGORITHM FOR MULTI-CHANNEL ACTIVE NOISE CONTROL AND VARIABLE SPATIAL SMOOTHING .................................................................................................................................................... 4695
Y. J. Chu, S. C. Chan, C. M. Mak, M. Wu

ADAPT-THEN-COMBINE FULL WAVEFORM INVERSION FOR DISTRIBUTED SUBSURFACE IMAGING IN SEISMIC NETWORKS .................................................................................................................................................... 4700
Ban-Sok Shin, Dmitriy Shutin

DATA FUSION FOR AUDIOVISUAL SPEAKER LOCALIZATION: EXTENDING DYNAMIC STREAM WEIGHTS TO THE SPATIAL DOMAIN .................................................................................................................................................... 4705
Julio Wissing, Benedikt Boenninghoff, Dorothea Kolossa, Tsubasa Ochiai, Marc Delcroix, Keisuke Kinoshita, Tomohiro Nakatani, Shoko Araki, Christopher Schymura

TOWARDS ROBUST TRAINING OF MULTI-SENSOR DATA FUSION NETWORK AGAINST ADVERSARIAL EXAMPLES IN SEMANTIC SEGMENTATION .................................................................................................................................................... 4710
Youngjoon Yu, Hong Joo Lee, Byeong Cheon Kim, Jung Uk Kim, Yong Man Ro

SPCOM-1: SIGNAL PROCESSING FOR NETWORKS

DATA-DRIVEN ADAPTIVE NETWORK RESOURCE SLICING FOR MULTI-TENANT NETWORKS .................................................................................................................................................... 4715
Navid Reyhanian, Hamid Farmanbar, Zhi-Quan Luo

DISTRIBUTED SCHEDULING USING GRAPH NEURAL NETWORKS .................................................................................................................................................... 4720
Zhongyuan Zhao, Gunjan Verma, Chirag Rao, Ananthram Swami, Santiago Segarra

EFFICIENT POWER ALLOCATION USING GRAPH NEURAL NETWORKS AND DEEP ALGORITHM UNFOLDING .................................................................................................................................................... 4725
Arindam Chowdhury, Gunjan Verma, Chirag Rao, Ananthram Swami, Santiago Segarra

A SAMPLE-EFFICIENT SCHEME FOR CHANNEL RESOURCE ALLOCATION IN NETWORKED ESTIMATION .................................................................................................................................................... 4730
Marcos M. Vasconcelos, Urbashi Mitra

AN EFFICIENT LINEAR PROGRAMMING ROUNDDING-AND-REFINEMENT ALGORITHM FOR LARGE-SCALE NETWORK SLICING PROBLEM .................................................................................................................................................... 4735
Wei-Kun Chen, Ya-Feng Liu, Yu-Hong Dai, Zhi-Quan Luo

EFFICIENT MIGRATION TO THE NEXT GENERATION OF NETWORKS BASED ON DIGITAL ANNEALING .................................................................................................................................................... 4740
Mohammad Javad-Kalbasi, Shahrokh Valaee

SPCOM-2: INFORMATION THEORY, CODING AND SECURITY

A TECHNIQUE FOR OFDM SYMBOL SLICING .................................................................................................................................................... 4745
Ana Perez-Neira, Miguel A. Lagunas

COMMUNICATION OVER BLOCK FADING CHANNELS – AN ALGORITHMIC PERSPECTIVE ON OPTIMAL TRANSMISSION SCHEMES .................................................................................................................................................... 4750
Holger Boche, Rafael F. Schaefer, H. Vincent Poor
SECURE UAV COMMUNICATIONS UNDER UNCERTAIN EAVESDROPPERS LOCATIONS .......... 4755
Silei Wang, Fanxiang Kong, Qiang Li

ON STRATEGIC JAMMING IN DISTRIBUTED DETECTION NETWORKS ........................................ 4760
Chen Quan, Baoccheng Geng, Pramod K. Varshney

REAL NUMBER SIGNAL PROCESSING CAN DETECT DENIAL-OF-SERVICE ATTACKS .......... 4765
Holger Boche, Rafael F. Schaefer, H. Vincent Poor

A HYBRID APPROACH TO CODED COMPRESSED SENSING WHERE COUPLING TAKES
PLACE VIA THE OUTER CODE ................................................................................................................ 4770
Jamison R. Ebert, Vamsi K. Amalladinne, Jean-Francois Chamberland, Krishna R. Narayanan

SPCOM-3: BEAMFORMING

GLOBALLY OPTIMAL BEAMFORMING FOR RATE SPLITTING MULTIPLE ACCESS ............... 4775
Bho Matthiesen, Yijie Mao, Petar Popovski, Bruno Clerckx

BEAM FOCUSING FOR MULTI-USER MIMO COMMUNICATIONS WITH DYNAMIC
METASURFACE ANTENNAS .................................................................................................................... 4780
Haiyang Zhang, Nir Shlezinger, Francesco Guidi, Davide Dardari, Mohammadreza F. Imani,
Yonina C. Eldar

PUSHING THE LIMIT OF TYPE I CODEBOOK FOR FDD MASSIVE MIMO
BEAMFORMING: A CHANNEL COVARIANCE RECONSTRUCTION APPROACH .................. 4785
Kai Li, Ying Li, Lei Cheng, Qingjiang Shi, Zhi-Quan Luo

FIRST-ORDER FAST ALGORITHM FOR STRUCTURALLY OPTIMAL MULTI-GROUP
MULTICAST BEAMFORMING IN LARGE-SCALE SYSTEMS ............................................................ 4790
Chong Zhang, Min Dong, Ben Liang

ANALOG BEAMFORMING WITH ANTENNA SELECTION FOR LARGE-SCALE
ANTENNA ARRAYS ............................................................................................................................ 4795
Aakash Arora, Christos G. Tsinos, Bhavani Shankar Mysore R, Symeon Chatzinotas, Björn
Ottersten

BEAMFORMING FOR BIDIRECTIONAL MIMO FULL DUPLEX UNDER THE JOINT SUM
POWER AND PER ANTENNA POWER CONSTRAINTS ................................................................. 4800
Chandan Kumar Sheemar, Dirk Slock

SPCOM-4: CHANNEL ESTIMATION FOR MIMO AND MULTIUSER SYSTEMS

ITERATIVE REWEIGHTED ALGORITHMS FOR JOINT USER IDENTIFICATION AND
CHANNEL ESTIMATION IN SPATIALLY CORRELATED MASSIVE MTC ........................................ 4805
Hamza Djelouat, Markus Leinonen, Markku Juntti

MILLIMETER WAVE MIMO CHANNEL ESTIMATION WITH 1-BIT SPATIAL SIGMA-
DELTA ANALOG-TO-DIGITAL CONVERTERS ............................................................................. 4810
R. S. Prasobh Sankar, Sundeep Prabhakar Chepuri

AN EFFICIENT ALGORITHM FOR DEVICE DETECTION AND CHANNEL ESTIMATION
IN ASYNCHRONOUS IOT SYSTEMS ................................................................................................. 4815
Liang Liu, Ya-Feng Liu
KALMAN FILTER BASED MIMO CSI PHASE RECOVERY FOR COTS WIFI DEVICES ........................................ 4820
Chu Li, Jeremy Brauer, Aydin Sezgin, Christian Zenger

IMPROVED ATOMIC NORM BASED CHANNEL ESTIMATION FOR TIME-VARYING NARROWBAND LEAKED CHANNELS ........................................................................................................ 4825
Jianxiu Li, Urbashi Mitra

BAYESIAN MASSIVE MIMO CHANNEL ESTIMATION WITH PARAMETER ESTIMATION USING LOW-RESOLUTION ADCS ........................................................................................................ 4830
Shuai Huang, Deqiang Qiu, Trac D. Tran

SPCOM-5: DETECTION AND DECODING

OPTIMAL DETECTION IN THE PRESENCE OF NON-GAUSSIAN JAMMING .................................................. 4835
Khalid A. Almahorg, Ramy H. Gohary

AN EFFICIENT ACTIVE SET ALGORITHM FOR COVARIANCE BASED JOINT DATA AND ACTIVITY DETECTION FOR MASSIVE RANDOM ACCESS WITH MASSIVE MIMO ...................................................... 4840
Ziyou Wang, Zhilin Chen, Ya-Feng Liu, Foad Sohrab, Wei Yu

NEURAL LAYERED MIN-SUM DECODING FOR PROTOGRAPH LDPC CODES .............................................. 4845
Dexin Zhang, Jincheng Dai, Kailin Tan, Kai Niu, Mingzhe Chen, H. Vincent Poor, Shuguang Cui

INTEGER CARRIER FREQUENCY OFFSET ESTIMATION IN OFDM WITH ZADOFF-CHU SEQUENCES ................................................................................................................................. 4850
John D. Roth, David A. Garren, R. Clark Robertson

PLUG-AND-PLAY LEARNED GAUSSIAN-MIXTURE APPROXIMATE MESSAGE PASSING .................................. 4855
Osman Musa, Peter Jung, Giuseppe Caire

LOW-LATENCY POLAR DECODER USING OVERLAPPED SCL PROCESSING .................................................. 4860
Dongyun Kam, Byeong Yong Kong, Youngjoo Lee

SPCOM-6: SYSTEM DESIGN AND OPTIMIZATION

MODULAR BINARY TREE ARCHITECTURE FOR DISTRIBUTED LARGE INTELLIGENT SURFACE ......................................................................................................................... 4865
Juan Vidal Alegría, Fredrik Rusek, Jesús Rodríguez Sánchez, Ove Edfors

STOCHASTIC SUCCESSIVE WEIGHTED SUM-RATE MAXIMIZATION FOR MULTIUSER MIMO SYSTEMS WITH FINITE-ALPHABET INPUTS ......................................................................................... 4870
Xin Guan, Xiaotong Zhao, Qingjiang Shi

RATE 1 QUASI ORTHOGONAL UNIVERSAL TRANSMISSION AND COMBINING FOR MIMO SYSTEMS ACHIEVING FULL DIVERSITY .................................................................................................. 4875
Barak Avraham, Uri Erez, Elad Domanovitz

ENERGY EFFICIENCY OPTIMIZATION TECHNIQUE FOR SWIPT-ENABLED MULTI-GROUP MULTICASTING SYSTEMS WITH HETEROGENEOUS USERS ........................................................................ 4880
Sumit Gautam, Symeon Chatzinotas, Björn Ottersten

MULTI-BRANCH TOMLINSON-HARASHIMA PRECODING FOR RATE SPLITTING BASED SYSTEMS WITH MULTIPLE ANTENNAS ................................................................................................. 4885
A. R. Flores, R. C. De Lamare, B. Clerckx
DIVIDE AND CONQUER: ONE-BIT MIMO-OFDM DETECTION BY INEXACT EXPECTATION MAXIMIZATION ………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………
ON INFORMATION ASYMMETRY IN ONLINE REINFORCEMENT LEARNING................................. 4955
Ezra Tampubolon, Haris Ceribasic, Holger Boche

JAMMING STRATEGY GENERATION FOR HIDDEN COMMUNICATION MODES VIA
GRAPH CONVOLUTION NETWORKS ....................................................................................... 4960
Fanxiang Kong, Qiang Li, Huaizong Shao

CONTRASTIVE SELF-SUPERVISED LEARNING FOR WIRELESS POWER CONTROL.............. 4965
Navid Naderializadeh

SPTM-1: DETECTION THEORY AND METHODS 1

MEASURE-TRANSFORMED COVARIANCE TEST FOR ROBUST SPECTRUM SENSING .......... 4970
Yair Sorek, Koby Todros

SEARCHING FOR ANOMALIES WITH MULTIPLE PLAYS UNDER DELAY AND
SWITCHING COSTS ............................................................................................................... 4975
Tidhar Lambez, Kobi Cohen

ROBUST ESTIMATION OF HIGH-ORDER PHASE DYNAMICS USING VARIATIONAL
BAYES INFERENC E.................................................................................................................. 4980
Fabio Fabozzi, Stéphanie Bidon, Sébastien Roche

ROBUST PCA THROUGH MAXIMUM CORRENTROPY POWER ITERATIONS ...................... 4985
Jean P. Chereau, Bruno Scalzo, Danilo P. Mandic

SCORE-BASED CHANGE DETECTION FOR GRADIENT-BASED LEARNING MACHINES.... 4990
Lang Liu, Joseph Salmon, Zaid Harchaoui

SUPER-RESOLUTION OF PERIODIC SIGNALS FROM SHORT SEQUENCES OF SAMPLES ...... 4995
Marek W. Rupniewski

SPTM-2: DETECTION THEORY AND METHODS 2

QUICKEST CHANGE DETECTION WITH TIME INCONSISTENT ANTICIPATORY
AGENTS IN CYBER-PHYSICAL SYSTEMS.............................................................................. 5000
Vikram Krishnamurthy

TREATMENT EFFECT ESTIMATION USING INARIANT RISK MINIMIZATION ...................... 5005
Abhin Shah, Kartik Ahuja, Karthikeyan Shanmugam, Dennis Wei, Kush R. Varshney, Amit
Dhurandhar

AN F-TEST FOR POLYNOMIAL FREQUENCY MODULATION .................................................. 5010
Kian Blanchette, Wesley Burr, Glen Takahara

QUICKEST JOINT DETECTION AND CLASSIFICATION OF FAULTS IN STATISTI CALLY
PERIODIC PROCESSES........................................................................................................... 5015
Taposh Banerjee, Smruti Padhy, Ahmad Taha, Eugene John

AN ASYMPTOTICALLY POINTWISE OPTIMAL PROCEDURE FOR SEQUENTIAL JOINT
DETECTION AND ESTIMATION............................................................................................ 5020
Dominik Reinhard, Michael Fauß, Abdelhak M. Zoubir
LOCALLY OPTIMAL DETECTION OF STOCHASTIC TARGETED UNIVERSAL ADVERSARIAL PERTURBATIONS .......................................................................................................................... 5025
  Amish Goel, Pierre Moulin

SPTM-3: ESTIMATION, DETECTION AND LEARNING OVER NETWORKS 1

A DECENTRALIZED VARIANCE-REDUCED METHOD FOR STOCHASTIC OPTIMIZATION OVER DIRECTED GRAPHS .......................................................................................................................... 5030
  Muhammad I. Qureshi, Ran Xin, Soummya Kar, Usman A. Khan

ON DISTRIBUTED COMPOSITE TESTS WITH DEPENDENT OBSERVATIONS IN WSN ........................................ 5035
  Juan Augusto Maya, Leonardo Rey Vega

BYZANTINE-RESILIENT DECENTRALIZED TD LEARNING WITH LINEAR FUNCTION APPROXIMATION .......................................................................................................................... 5040
  Zhaoxian Wu, Han Shen, Tianyi Chen, Qing Ling

ON THE EFFECT OF SPATIAL CORRELATION ON DISTRIBUTED ENERGY DETECTION OF A STOCHASTIC PROCESS ................................................................................................................... 5045
  Juan Augusto Maya, Leonardo Rey Vega

SPTM-4: ESTIMATION, DETECTION AND LEARNING OVER NETWORKS 2

PROVABLY FAST ASYNCHRONOUS AND DISTRIBUTED ALGORITHMS FOR PAGERANK CENTRALITY COMPUTATION .......................................................................................................................... 5050
  Yiran He, Hoi-To Wai

DECENTRALIZED OPTIMIZATION OVER NOISY, RATE-CONSTRAINED NETWORKS: HOW WE AGREE BY TALKING ABOUT HOW WE DISAGREE .......................................................................................................................... 5055
  Rajarshi Saha, Stefano Rini, Milind Rao, Andrea Goldsmith

A MULTIPLE ACCESS CHANNEL GAME USING LATENCY METRIC ........................................................................ 5060
  Audrey Garnaev, Athina Petropulu, Wade Trappe

SPTM-5: SAMPLING, MULTIRATE SIGNAL PROCESSING AND DIGITAL SIGNAL PROCESSING 1

LINEAR COMPUTATION CODING .......................................................................................................................... 5065
  Ralf R. Müller, Bernhard Gäde, Ali Bereyhi

SPECTRAL FOLDING AND TWO-CHANNEL FILTER-BANKS ON ARBITRARY GRAPHS ........................................ 5070
  Eduardo Pavez, Benjamin Girault, Antonio Ortega, Philip A. Chou

SPARSE TIME-FREQUENCY REPRESENTATION VIA ATOMIC NORM MINIMIZATION ........................................ 5075
  Tsubasa Kusano, Kohei Yatabe, Yasuhiro Oikawa

MESSAGE TRANSMISSION OVER RAPIDLY TIME-VARYING CHANNELS .......................................................... 5080
  Alihan Kaplan, Volker Pohl

A ROBUST AND EFFICIENT MULTI-SCALE SEASONAL-TREND DECOMPOSITION ........................................ 5085
  Linxiao Yang, Qingsong Wen, Bo Yang, Liang Sun

NOISE-ASSISTED MULTIVARIATE VARIATIONAL MODE DECOMPOSITION .......................................................... 5090
  Charilaos A. Zisou, Georgios K. Apostolidis, Leontios J. Hadjileontiadis
SPTM-6: SAMPLING, MULTIRATE SIGNAL PROCESSING AND DIGITAL SIGNAL PROCESSING 2

APPROXIMATE WEIGHTED C R CODED MATRIX MULTIPLICATION ................................................................. 5095
Neophytos Charalambides, Mert Pilanci, Alfred O. Hero

PERIODIC SIGNAL DENOISING: AN ANALYSIS-SYNTHESIS FRAMEWORK BASED ON Ramanujan Filter Banks and Dictionaries ................................................................................................. 5100
Pranav Kulkarni, P. P. Vaidyanathan

COMPRESSIVE SIGNAL RECOVERY UNDER SENSING MATRIX ERRORS COMBINED WITH UNKNOWN MEASUREMENT GAINS ........................................................................................................... 5105
Jian Vora, Ajit Rajwade

GRID OPTIMIZATION FOR MATRIX-BASED SOURCE LOCALIZATION UNDER INHOMOGENEOUS SENSOR TOPOLOGY .................................................................................................................. 5110
Hao Sun, Junting Chen

MSR-GAN: MULTI-SEGMENT RECONSTRUCTION VIA ADVERSARIAL LEARNING ........................................ 5115
Mona Zehni, Zhizhen Zhao

COUNT SKETCH WITH ZERO CHECKING: EFFICIENT RECOVERY OF HEAVY COMPONENTS .......................................................... 5120
Guanqiang Zhou, Zhi Tian

SPTM-7: ESTIMATION THEORY AND METHODS 1

NUMERICAL SOLUTION OF STOCHASTIC DIFFERENTIAL EQUATIONS IN STIEFEL MANIFOLDS VIA TANGENT SPACE PARAMETRIZATION ................................................................. 5125
Zhichao Wang, Victor Solo

ON THE ACCURACY LIMIT OF JOINT TIME-DELAY/DOPPLER/ACCELERATION ESTIMATION WITH A BAND-LIMITED SIGNAL ....................................................................................... 5130
H. McPhee, L. Ortega, J. Vilà-Valls, E. Chaumette

AUTOMATIC ORDER SELECTION IN AUTOREGRESSIVE MODELING WITH APPLICATION IN EEG SLEEP-STAGE CLASSIFICATION .................................................................................................. 5135
Farah Nassif, Soosan Beheshti

NEW VARIANTS OF DFA BASED ON LOESS AND LOWESS METHODS: GENERALIZATION OF THE DETRENDING MOVING AVERAGE .......................................................... 5140
Bastien Berthelot, Eric Grivel, Pierrick Legrand

PARAMETER ESTIMATION FOR STUDENT’S T VAR MODEL WITH MISSING DATA .......................... 5145
Rui Zhou, Junyan Liu, Sandeep Kumar, Daniel P. Palomar

FAST AND ROBUST ADMM FOR BLIND SUPER-RESOLUTION ................................................................ 5150
Yifan Ran, Wei Dai

SPTM-8: ESTIMATION THEORY AND METHODS 2

NONSTATIONARY PORTFOLIOS: DIVERSIFICATION IN THE SPECTRAL DOMAIN .................................. 5155
Bruno Scalzo, Alvaro Arroyo, Ljubiša Stankovic, Danilo P. Mandic
A TYLER-TYPE ESTIMATOR OF LOCATION AND SCATTER LEVERAGING Riemannian Optimization .......................................................... 5160
  A. Collas, F. Bouchard, A. Breloy, C. Ren, G. Ginolhac, J.-P. Ovarlez

STATISTICAL PROPERTIES OF A MODIFIED WELCH METHOD THAT USES SAMPLE PERCENTILES ................................................................. 5165
  Felix Schwock, Shima Abadi

SWITCHED HAWKES PROCESSES .................................................................................................................. 5170
  Namrata Nadagouda, Mark A. Davenport

AN ADAPTIVE REGULARIZATION APPROACH TO PORTFOLIO OPTIMIZATION .................................................. 5175
  Tarig Ballal, Abdelrahman S. Abdelrahman, Ali H. Muaibbel, Tareq Y. Al-Naffouri

ACTIVE ESTIMATION FROM MULTIMODAL DATA .................................................................................. 5180
  Arpan Mukherjee, Ali Tajer, Pin-Yu Chen, Payel Das

SPTM-9: ESTIMATION, DETECTION AND LEARNING OVER NETWORKS 3

NETWORK CLASSIFIERS BASED ON SOCIAL LEARNING ...................................................................................... 5185
  Virginia Bordignon, Stefan Vlaski, Vincenzo Matta, Ali H. Sayed

BAYES-OPTIMAL METHODS FOR FINDING THE SOURCE OF A CASCADE ........................................................... 5190
  Anirudh Sridhar, H. Vincent Poor

PRIVATE WIRELESS FEDERATED LEARNING WITH ANONYMOUS OVER-THE-AIR COMPUTATION .......................................................... 5195
  Burak Hasircioglu, Deniz Gündüz

SCALABLE MULTILEVEL QUANTIZATION FOR DISTRIBUTED DETECTION ...................................................... 5200
  Gökhan Gül, Michael Baßler

STABILITY OF ALGEBRAIC NEURAL NETWORKS TO SMALL PERTURBATIONS ............................................. 5205
  Alejandro Parada-Mayorga, Alejandro Ribeiro

RESOLUTION LIMITS OF 20 QUESTIONS SEARCH STRATEGIES FOR MOVING TARGETS .................................................. 5210
  Lin Zhou, Alfred Hero

SPTM-10: DISTRIBUTED LEARNING OVER GRAPHS

GRAMIAN-BASED ADAPTIVE COMBINATION POLICIES FOR DIFFUSION LEARNING OVER NETWORKS .................................................................... 5215
  Y. Efe Erginbas, Stefan Vlaski, Ali H. Sayed

GRAPH-ADAPTIVE INCREMENTAL LEARNING USING AN ENSEMBLE OF GAUSSIAN PROCESS EXPERTS .......................................................... 5220
  Konstantinos D. Polyzos, Qin Lu, Georgios B. Giannakis

FAST DECENTRALIZED LINEAR FUNCTIONS VIA SUCCESSIVE GRAPH SHIFT OPERATORS ............................................................... 5225
  Siavash Mollaebrahim, Daniel Romero, Baltasar Beferull-Lozano

ONLINE LEARNING OF TIME-VARYING SIGNALS AND GRAPHS ............................................................. 5230
  Stefania Sardellitti, Sergio Barbarossa, Paolo Di Lorenzo
KERNEL REGRESSION ON GRAPHS IN RANDOM FOURIER FEATURES SPACE .............................................. 5235
Vitor R. M. Elias, Vinay C. Gogineni, Wallace A. Martins, Stefan Werner

GRAPH-HOMOMORPHIC PERTURBATIONS FOR PRIVATE DECENTRALIZED LEARNING .................................................. 5240
Stefan Vlaski, Ali H. Sayed

SPTM-11: GRAPHS NEURAL NETWORKS

VARIANCE-CONSTRAINED LEARNING FOR STOCHASTIC GRAPH NEURAL NETWORKS ...................................................... 5245
Zhan Gao, Elvin Isufi, Alejandro Ribeiro

GRAPH NEURAL NETWORK FOR LARGE-SCALE NETWORK LOCALIZATION ......................................................... 5250
Wenzhong Yan, Di Jin, Zhidi Lin, Feng Yin

GRAPHON AND GRAPH NEURAL NETWORK STABILITY ................................................................................. 5255
Luana Ruiz, Zhiyang Wang, Alejandro Ribeiro

GRAPH NEURAL NETWORKS FOR DECENTRALIZED CONTROLLERS ................................................................. 5260
Fernando Gama, Ekaterina Tolstaya, Alejandro Ribeiro

NONLINEAR STATE-SPACE GENERALIZATIONS OF GRAPH CONVOLUTIONAL NEURAL NETWORKS .................. 5265
Luana Ruiz, Fernando Gama, Alejandro Ribeiro, Elvin Isufi

WIDE AND DEEP GRAPH NEURAL NETWORKS WITH DISTRIBUTED ONLINE LEARNING ...................................................... 5270
Zhan Gao, Alejandro Ribeiro, Fernando Gama

SPTM-12: SAMPLING, FILTERING AND DENOISING OVER GRAPHS

DESIGN OF GRAPH SIGNAL SAMPLING MATRICES FOR ARBITRARY SIGNAL SUBSPACES .................................................. 5275
Junya Hara, Koki Yamada, Shunsuke Ono, Yuichi Tanaka

GRAPH SIGNAL DENOISING USING NESTED-STRUCTURED DEEP ALGORITHM UNROLLING .................................................. 5280
Masatoshi Nagahama, Koki Yamada, Yuichi Tanaka, Stanley H. Chan, Yonina C. Eldar

IDENTIFYING FIRST-ORDER LOWPASS GRAPH SIGNALS USING PERRON FROBENIUS THEOREM ......................................... 5285
Yiran He, Hoi-To Wai

GRAPH SIGNAL DENOISING VIA UNROLLING NETWORKS ...................................................................................... 5290
Siheng Chen, Yonina C. Eldar

ADAPTIVE SUBSAMPLING OF MULTIDOMAIN SIGNALS WITH PRODUCT GRAPHS .......................................................... 5295
Théo Gnassounou, Pierre Humbert, Laurent Oudre

ROBUST GRAPH-FILTER IDENTIFICATION WITH GRAPH DENOISING REGULARIZATION .............................................. 5300
Samuel Rey, Antonio G. Marques
SPTM-13: MODELS, METHODS AND ALGORITHMS 1

FAST AND PROVABLE ROBUST PCA VIA NORMALIZED COHERENCE PURSUIT
Mostafa Rahmani, Ping Li

ALIGNING SETS OF TEMPORAL SIGNALS WITH RIEMANNIAN GEOMETRY AND KOOPMAN OPERATOR
Ohad Rahamim, Ronen Talmon

WEIGHT IDENTIFICATION THROUGH GLOBAL OPTIMIZATION IN A NEW HYSTERETIC NEURAL NETWORK MODEL
Elie Leroy, Arthur Marmin, Marc Castella, Laurent Duval

MULTIVIEW VARIATIONAL GRAPH AUTOENCODERS FOR CANONICAL CORRELATION ANALYSIS
Yacouba Kaloga, Pierre Borgnat, Sundeep Prabhakar Chepuri, Patrice Abry, Amaury Habrard

COGNITIVE MEMORY CONSTRAINED HUMAN DECISION MAKING BASED ON MULTI-SOURCE INFORMATION
Baocheng Geng, Quan Chen, Pramod K. Varshney

BINARY CONTROL AND DIGITAL-TO-ANALOG CONVERSION USING COMPOSITE NUV PRIORS AND ITERATIVE GAUSSIAN MESSAGE PASSING
Raphael Keusch, Hampus Malmberg, Hans-Andrea Loeliger

SPTM-14: MODELS, METHODS AND ALGORITHMS 2

OUTLIER-ROBUST KERNEL HIERARCHICAL-OPTIMIZATION RLS ON A BUDGET WITH AFFINE CONSTRAINTS
Konstantinos Slavakis, Masahiro Yukawa

ADAPTIVE REAL-TIME FILTER FOR PARTIALLY-OBSERVED BOOLEAN DYNAMICAL SYSTEMS
Mahdi Imani, Seyede Fatemeh Ghoreishi

IMPROVING THE ENERGY-EFFICIENCY OF A KALMAN FILTER USING UNRELIABLE MEMORIES
Jonathan Kern, Elsa Dupraz, Abdeldjalil Aïssa-El-Bey, François Leduc-Primeau

PARALLEL ITERATED EXTENDED AND SIGMA-POINT KALMAN SMOOTHERS
Fatemeh Yaghoobi, Adrien Corenflos, Sakira Hassan, Simo Särkkä

WIENER FILTER ON MEET/JOIN LATTICES
Bastian Seifert, Chris Wendler, Markus Püschel

SPTM-15: GRAPH TOPOLOGY INFERENCE AND CLUSTERING

LEARNING BOLLOBÁS-RIORDAN GRAPHS UNDER PARTIAL OBSERVABILITY
Michele Cirillo, Vincenzo Matta, Ali H. Sayed

LEARNING SPARSE GRAPH LAPLACIAN WITH K EIGENVECTOR PRIOR VIA ITERATIVE GLASSO AND PROJECTION
Saghar Bagheri, Gene Cheung, Antonio Ortega, Fen Wang
LEARNING MIXED MEMBERSHIP FROM ADJACENCY GRAPH VIA SYSTEMATIC EDGE QUERY: IDENTIFIABILITY AND ALGORITHM ................................................................. 5370
  Shahana Ibrahim, Xiao Fu

CONVERGENCE ANALYSIS OF THE GRAPH-TOPOLOGY-INFERENC KERNEL LMS ALGORITHM ........................................................................................................... 5375
  Mircea Moscu, Ricardo Borsoi, Cédric Richard

AN EFFICIENT ALTERNATING DIRECTION METHOD FOR GRAPH LEARNING FROM SMOOTH SIGNALS ......................................................................................... 5380
  Xiaolu Wang, Chaorui Yao, Haoyu Lei, Anthony Man-Cho So

TOPOLOGICAL VOLterra FILTERS ........................................................................................................ 5385
  Geert Leus, Maosheng Yang, Mario Coutino, Elvin Isufi

SPTM-16: GRAPH TOPOLOGY INFERENCE

NETWORK TOPOLOGY INFERENCE WITH GRAPHON SPECTRAL PENALTIES .................. 5390
  T. Mitchell Roddenberry, Madeline Navarro, Santiago Segarra

NETWORK TOPOLOGY CHANGE-POINT DETECTION FROM GRAPH SIGNALS WITH PRIOR SPECTRAL SIGNATURES............................................................... 5395
  Chiraag Kaushik, T. Mitchell Roddenberry, Santiago Segarra

ONLINE TIME-VARYING TOPOLOGY IDENTIFICATION VIA PREDICTION-CORRECTION ALGORITHMS ................................................................. 5400
  Alberto Natali, Mario Coutino, Elvin Isufi, Geert Leus

GRAPH LEARNING UNDER SPECTRAL SPARSITY CONSTRAINTS ........................................ 5405
  B. Subbareddy, Aditya Siripuram, Jingxin Zhang

A GRAPH LEARNING ALGORITHM BASED ON GAUSSIAN MARKOV RANDOM FIELDS AND MINIMAX CONCAVE PENALTY ......................................................... 5410
  Tatsuya Koyakumaru, Masahiro Yukawa, Eduardo Pavez, Antonio Ortega

FIGLEARN: FILTER AND GRAPH LEARNING USING OPTIMAL TRANSPORT .................. 5415
  Matthias Minder, Zahra Farsijani, Dhruti Shah, Mireille El Gheche, Pascal Frossard

SPTM-17: SAMPLING, MULTIRATE SIGNAL PROCESSING AND DIGITAL SIGNAL PROCESSING 3

CONSTRUCTION OF UNIT-NORM TIGHT FRAME BASED PRECONDITIONER FOR SPARSE CODING ........................................................................................................... 5420
  Huang Bai, Chuanrong Hong, Xiumei Li

SPARSE HIGH-ORDER PORTFOLIOS VIA PROXIMAL DCA AND SCA ................................ 5425
  Jinxin Wang, Zengde Deng, Taoli Zheng, Anthony Man-Cho So

A CONVEX PENALTY FOR BLOCK-SPARSE SIGNALS WITH UNKNOWN STRUCTURES ........ 5430
  Hiroki Kuroda, Daichi Kitahara, Akira Hirabayashi

EVENT-DRIVEN MODULO SAMPLING ......................................................................................... 5435
  Dorian Florescu, Felix Krahmer, Ayush Bhandari
NO RELAXATION: GUARANTEED RECOVERY OF FINITE-VALUED SIGNALS FROM UNDERSAMPLED MEASUREMENTS .......................................................... 5440
Pulak Sarangi, Piya Pal

ERROR ESTIMATES IN SECOND-ORDER CONTINUOUS-TIME SIGMA-DELTA MODULATORS .............................................................................................................................................. 5445
Dilshad Surroop, Pascal Combes, Philippe Martin

SPTM-18: SAMPLING THEORY, ANALYSIS AND METHODS

BANRAW: BAND-LIMITED RADAR WAVEFORM DESIGN VIA PHASE RETRIEVAL ......................................................... 5449
Samuel Pinilla, Kumar Vijay Mishra, Brian M. Sadler, Henry Arguello

SUB-NYQUIST MULTICHANNEL BLIND DECONVOLUTION .............................................................................................................. 5454
Satish Mulleti, Kiryung Lee, Yonina C. Eldar

MODIFIED ARC SINELAW FOR ONE-BIT SAMPLED STATIONARY SIGNALS WITH TIME-VARYING THRESHOLDS ........................................................................................................................................... 5459
Arian Eamaz, Farhang Yeganegi, Mojtaba Soltanalian

NEAR-OPTIMAL RESAMPLING IN PARTICLE FILTERS USING THE ISING ENERGY MODEL .......................................................................................................................................................................................... 5464
Muhammed Tahsin Rahman, Mohammad Javad-Kalbasi, Shahrokh Valaee

TIME-DOMAIN CONCENTRATION AND APPROXIMATION OF COMPUTABLE BANDLIMITED SIGNALS ......................................................................................................................................................... 5469
Holger Boche, Ullrich J. Mönich

GUARANTEED RECONSTRUCTION FROM INTEGRATE-AND-FIRE NEURONS WITH ALPHA SYNAPTIC ACTIVATION ............................................................................................................................................. 5474
Marek Hilton, Roxana Alexandru, Pier Luigi Dragotti

SPTM-19: INFERENCE OVER GRAPHS

SOCIAL LEARNING UNDER INFERENTIAL ATTACKS ................................................................................................................. 5479
Konstantinos Ntemos, Virginia Bordignon, Stefan Vlaski, Ali H. Sayed

SEGREGATION IN SOCIAL NETWORKS: MARKOV BRIDGE MODELS AND ESTIMATION ......................................................................................................................................................................................... 5484
Vikram Krishnamurthy, Rui Luo, Buddhika Nettasinghe

CONTROLLED TESTING AND ISOLATION FOR SUPPRESSING COVID-19 ..................................................................................... 5489
Kobi Cohen, Amir Leshem

TWO-STAGE GRAPH-CONSTRAINED GROUP TESTING: THEORY AND APPLICATION ............................................................................................ 5494
Saurabh Sihag, Ali Tajer, Urbashi Mitra

UNVEILING ANOMALOUS NODES VIA RANDOM SAMPLING AND CONSENSUS ON GRAPHS ................................................................................................................................. 5499
Vassilis N. Ioannidis, Dimitris Berberidis, Georgios B. Giannakis

ESTIMATING FIEDLER VALUE ON LARGE NETWORKS BASED ON RANDOM WALK OBSERVATIONS ............................................................................................................................... 5504
Alexandre Reiffers-Masson, Thierry Chonavel, Yezekael Hayel
## SPTM-20: SIGNAL PROCESSING OVER GRAPHS AND SPARSITY-AWARE SIGNAL PROCESSING

**Orthogonality and Zero DC Tradeoffs in Biorthogonal Graph Filterbanks**

*Dion E. O. Tzamarias, Eduardo Pavez, Benjamin Girault, Antonio Ortega, Ian Blanes, Joan Serra-Sagristà*

**Graph Signal Compression via Task-Based Quantization**

*Pei Li, Nir Shlezinger, Haiyang Zhang, Baoyun Wang, Yonina C. Eldar*

**A Partially Collapsed Gibbs Sampler for Unsupervised Nonnegative Sparse Signal Restoration**

*M. C. Amrouche, H. Carfantan, J. Idier*

**A Structure-Guided and Sparse-Representation-Based 3D Seismic Inversion Method**

*Bin She, Yaojun Wang, Guangmin Hu*

## SPTM-21: OPTIMIZATION METHODS FOR SIGNAL PROCESSING

**Accelerating Frank-Wolfe with Weighted Average Gradients**

*Yilang Zhang, Bingcong Li, Georgios B. Giannakis*

**YAPA: Accelerated Proximal Algorithm for Convex Composite Problems**

*Giovanni Chierchia, Mireille El Gheche*

**Data Discovery Using Lossless Compression-Based Sparse Representation**

*Elyas Sabeti, Peter X. K. Song, Alfred O. Hero*

**Safe Screening for Sparse Regression with the Kullback-Leibler Divergence**

*Cássio F. Dantas, Emmanuel Soubies, Cédric Févotte*

**On the Convergence of Randomized Bregman Coordinate Descent for Non-Lipschitz Composite Problems**

*Tianxiang Gao, Songtao Lu, Jia Liu, Chris Chu*

**A Global Cayley Parametrization of Stiefel Manifold for Direct Utilization of Optimization Mechanisms Over Vector Spaces**

*Keita Kume, Isao Yamada*

## SPTM-22: SIGNAL PROCESSING THEORY AND METHODS

**Training Logical Neural Networks by Primal–Dual Methods for Neuro-Symbolic Reasoning**

*Songtao Lu, Naweed Khan, Ismail Yunus Akhalwaya, Ryan Riegel, Lior Horesh, Alexander Gray*
**SPTM-23: BAYESIAN SIGNAL PROCESSING**

COOPERATIVE PARAMETER TRACKING ON THE UNIT SPHERE USING DISTRIBUTED ADAPT-THEN-COMBINE PARTICLE FILTERS AND PARALLEL TRANSPORT ........................................... 5564

Caio G. De Figueredo, Claudio J. Bordin, Marcelo G. S. Bruno

BAYESIAN ESTIMATION OF A TAIL-INDEX WITH MARGINALIZED THRESHOLD ................................... 5569

Douglas E. Johnston, Petar M. Djuric

BLOCK KALMAN FILTER: AN ASYMPTOTIC BLOCK PARTICLE FILTER IN THE LINEAR GAUSSIAN CASE ......................................................................................................................... 5574

Rui Min, Christelle Garnier, François Septier, John Klein

PARTICLE GIBBS SAMPLING FOR REGIME-SWITCHING STATE-SPACE MODELS ................................. 5579

Yousef El-Laham, Liu Yang, Heather J. Lynch, Petar M. Djuric, Mónica F. Bugallo

ADAPTIVE IMPORTANCE SAMPLING VIA AUTO-REGRESSIVE GENERATIVE MODELS AND GAUSSIAN PROCESSES .............................................................................................................. 5584

Hechuan Wang, Mónica F. Bugallo, Petar M. Djuric

VARIATIONAL PARAMETER LEARNING IN SEQUENTIAL STATE-SPACE MODEL VIA PARTICLE FILTERING .............................................................. 5589

Chenhao Li, Simon J. Godsill

**SPTM-24: SPARSITY-AWARE PROCESSING**

CORRELATION-BASED ROBUST LINEAR REGRESSION WITH ITERATIVE OUTLIER REMOVAL ............................................................... 5594

Jian Ding, Jianji Wang, Yue Zhang, Yuanjie Li, Nanning Zheng

ON THE OPTIMALITY OF BACKWARD REGRESSION: SPARSE RECOVERY AND SUBSET SELECTION ......................................................................................................................... 5599

Sebastian Ament, Carla Gomes

GENERAL TOTAL VARIATION REGULARIZED SPARSE BAYESIAN LEARNING FOR ROBUST BLOCK-SPARSE SIGNAL RECOVERY .............................................................................. 5604

Aditya Sant, Markus Leinonen, Bhaskar D. Rao

AUTOMATIC REGISTRATION AND CLUSTERING OF TIME SERIES ......................................................... 5609

Michael Weylandt, George Michailidis

LOW-RANK ON GRAPHS PLUS TEMPORALLY SMOOTH SPARSE DECOMPOSITION FOR ANOMALY DETECTION IN SPATIOTEMPORAL DATA ........................................................................ 5614

Seyyid Emre Sofuoglu, Selin Aviyente

A PARALLEL ALGORITHM FOR PHASE RETRIEVAL WITH DICTIONARY LEARNING .............................................. 5619

Tianyi Liu, Andreas M. Tillmann, Yang Yang, Yonina C. Eldar, Marius Pesavento

**SPE-1: SPEECH RECOGNITION 1: NEURAL TRANSDUCER MODELS 1**

IMPROVING RNN TRANSDUCER MODELING FOR SMALL-FOOTPRINT KEYWORD SPOTTING ............................................................................................................................ 5624

Yao Tian, Haitao Yao, Meng Cai, Yaming Liu, Zejun Ma
CASCADED ENCODERS FOR UNIFYING STREAMING AND NON-STREAMING ASR ........................ 5629
  Arun Narayanan, Tara N. Sainath, Ruoming Pang, Jiahui Yu, Chung-Cheng Chiu, Rohit Prabhavalkar, Ehsan Variani, Trevor Strohman

A BETTER AND FASTER END-TO-END MODEL FOR STREAMING ASR .......................... 5634
  Bo Li, Anmol Gulati, Jiahui Yu, Tara N. Sainath, Chung-Cheng Chiu, Arun Narayanan, Shuo-Yin Chang, Ruoming Pang, Yanzhang He, James Qin, Wei Han, Qiao Liang, Yu Zhang, Trevor Strohman, Yonghui Wu

EFFICIENT KNOWLEDGE DISTILLATION FOR RNN-TRANSUDER MODELS .......................... 5639
  Sankaran Panchapagesan, Daniel S. Park, Chung-Cheng Chiu, Yuan Shangguan, Qiao Liang, Alexander Gruenstein

PHONEME BASED NEURAL TRANSUDER FOR LARGE VOCABULARY SPEECH RECOGNITION ................................................................. 5644
  Wei Zhou, Simon Berger, Ralf Schlüter, Hermann Ney

RNN-T BASED OPEN-VOCABULARY KEYWORD SPOTTING IN MANDARIN WITH MULTI-LEVEL DETECTION .......................................................... 5649
  Zuozhen Liu, Ta Li, Pengyuan Zhang

SPE-2: SPEECH RECOGNITION 2: NEURAL TRANSUDER MODELS 2

ADVANCING RNN TRANSUDER TECHNOLOGY FOR SPEECH RECOGNITION .................. 5654
  George Saon, Zoltán Tüske, Daniel Bolanos, Brian Kingsbury

LESS IS MORE: IMPROVED RNN-T DECODING USING LIMITED LABEL CONTEXT AND PATH MERGING .......................................................... 5659
  Rohit Prabhavalkar, Yanzhang He, David Rybach, Sean Campbell, Arun Narayanan, Trevor Strohman, Tara N. Sainath

SIMPLEFLAT: A SIMPLE WHOLE-NETWORK PRE-TRAINING APPROACH FOR RNN TRANSUDER-BASED END-TO-END SPEECH RECOGNITION .................. 5664
  Takafumi Moriya, Takanori Ashihara, Tomohiro Tanaka, Tsubasa Ochiai, Hiroshi Sato, Atsushi Ando, Yusuke Ijima, Ryo Masumura, Yusuke Shinohara

ECHO STATE SPEECH RECOGNITION ........................................................................ 5669
  Harsh Shrivastava, Ankush Garg, Yuan Cao, Yu Zhang, Tara Sainath

USING SYNTHETIC AUDIO TO IMPROVE THE RECOGNITION OF OUT-OF-VOCABULARY WORDS IN END-TO-END ASR SYSTEMS ............................................. 5674
  Xianrui Zheng, Yulan Liu, Deniz Gunceler, Daniel Willett

SPE-3: SPEECH SYNTHESIS 1: ARCHITECTURE

WAVE-TACOTRON: SPECTROGRAM-FREE END-TO-END TEXT-TO-SPEECH SYNTHESIS ................................................................. 5679
  Ron J. Weiss, Rj Skerry-Ryan, Eric Battenberg, Soroosh Mariooryad, Diederik P. Kingma

PATNET: A PHONEME-LEVEL AUTOREGRESSIVE TRANSFORMER NETWORK FOR SPEECH SYNTHESIS ........................................ 5684
  Shiming Wang, Zhenhua Ling, Ruibo Fu, Jiangyan Yi, Jianhua Tao
MULTI-RATE ATTENTION ARCHITECTURE FOR FAST STREAMABLE TEXT-TO-SPEECH SPECTRUM MODELING ................................................................. 5689
Qing He, Zhiping Xiu, Thilo Koehler, Jilong Wu

END-TO-END TEXT-TO-SPEECH USING LATENT DURATION BASED ON VQ-VAE ........................................ 5694
Yusuke Yasuda, Xin Wang, Junichi Yamagishd

LIGHTSPEECH: LIGHTWEIGHT AND FAST TEXT TO SPEECH WITH NEURAL ARCHITECTURE SEARCH ................................................................ 5699
Renqian Luo, Xu Tan, Rui Wang, Tao Qin, Jinzhu Li, Sheng Zhao, Enhong Chen, Tie-Yan Liu

A NEW HIGH QUALITY TRAJECTORY TILING BASED HYBRID TTS IN REAL TIME .......................... 5704
Feng-Long Xie, Xin-Hui Li, Wen-Chao Su, Li Lu, Frank K. Soong

SPE-4: SPEECH SYNTHESIS 2: CONTROLLABILITY

PARALLEL TACOTRON: NON-AUTOREGRESSIVE AND CONTROLLABLE TTS ........................................... 5709
Isaac Elias, Heiga Zen, Jonathan Shen, Yu Zhang, Ye Jia, Ron J. Weiss, Yonghui Wu

FCL-TACO2: TOWARDS FAST, CONTROLLABLE AND LIGHTWEIGHT TEXT-TO-SPEECH SYNTHESIS ................................................................. 5714
Disong Wang, Liqin Deng, Yang Zhang, Nianzu Zheng, Yu Ting Yeung, Xiao Chen, Xunying Liu, Helen Meng

PROSODIC CLUSTERING FOR PHONEME-LEVEL PROSODY CONTROL IN END-TO-END SPEECH SYNTHESIS ................................................................. 5719
Alexandra Vioni, Myrsini Christidou, Nikolaos Ellinas, Georgios Yamvouakis, Panos Kakoulidis, Taehoon Kim, June Sig Sung, Hyoungmin Park, Aimilios Chalamandaris, Pirros Tsiakoulis

IMPROVING NATURALNESS AND CONTROLLABILITY OF SEQUENCE-TO-SEQUENCE SPEECH SYNTHESIS BY LEARNING LOCAL PROSODY REPRESENTATIONS ........................................ 5724
Cheng Gong, Longbiao Wang, Zhenhua Ling, Shaotong Guo, Ju Zhang, Jianwu Dang

MULTI-SPEAKER EMOTIONAL SPEECH SYNTHESIS WITH FINE-GRAINED PROSODY MODELING ................................................................. 5729
Chunhui Lu, Xue Wen, Ruolan Liu, Xiao Chen

EMOTION CONTROLLABLE SPEECH SYNTHESIS USING EMOTION-UNLABELED DATASET WITH THE ASSISTANCE OF CROSS-DOMAIN SPEECH EMOTION RECOGNITION ........................................................................ 5734
Xiong Cai, Dongyang Dai, Zhiyong Wu, Xiang Li, Jingbei Li, Helen Meng

SPE-5: SPEECH ENHANCEMENT 1: SPEECH SEPARATION

DUAL-PATH MODELING FOR LONG RECORDING SPEECH SEPARATION IN MEETINGS .......... 5739
Chenda Li, Zhuo Chen, Yi Luo, Cong Han, Tianyan Zhou, Keisuke Kinoshita, Marc Delcroix, Shinji Watanabe, Yanmin Qian

TIME-DOMAIN LOSS MODULATION BASED ON OVERLAP RATIO FOR MONOURAL CONVERSATIONAL SPEAKER SEPARATION ........................................................................ 5744
Hassan Taherian, Deliang Wang
CONTINUOUS SPEECH SEPARATION WITH CONFORMER............................................................... 5749
  Sanyuan Chen, Yu Wu, Zhuo Chen, Jian Wu, Jinyu Li, Takuya Yoshioka, Chengyi Wang,
  Shujie Liu, Ming Zhou

A FLOW-BASED NEURAL NETWORK FOR TIME DOMAIN SPEECH ENHANCEMENT........... 5754
  Martin Strauss, Bernd Edler

SANDGLASSET: A LIGHT MULTI-GRANULARITY SELF-ATTENTIVE NETWORK FOR
TIME-DOMAIN SPEECH SEPARATION.................................................................................. 5759
  Max W. Y. Lam, Jun Wang, Dan Su, Dong Yu

TRANSAMSK: A COMPACT AND FAST SPEECH SEPARATION MODEL BASED ON
TRANSFORMER .................................................................................................................. 5764
  Zining Zhang, Bingsheng He, Zhenjie Zhang

SPE-6: SPEECH ENHANCEMENT 2: SPEECH SEPARATION AND DEREVERBERATION

ONE SHOT LEARNING FOR SPEECH SEPARATION ............................................................ 5769
  Yuan-Kuei Wu, Kuan-Po Huang, Yu Tsao, Hung-Yi Lee

TRAINING NOISY SINGLE-CHANNEL SPEECH SEPARATION WITH NOISY ORACLE
SOURCES: A LARGE GAP AND A SMALL STEP................................................................. 5774
  Matthew Maciejewski, Jing Shi, Shinji Watanabe, Sanjeev Khudanpur

SPEAKER AND DIRECTION INFERRED DUAL-CHANNEL SPEECH SEPARATION ............... 5779
  Chenxing Li, Jiaming Xu, Nima Mesgarani, Bo Xu

SPEECH DEREVERBERATION USING VARIATIONAL AUTOENCODERS ................................ 5784
  Deepak Baby, Hervé Bourlard

REAL-TIME DENOISING AND DEREVERBERATION WTIH TINY RECURRENT U-NET .......... 5789
  Hyeong-Seok Choi, Sungjin Park, Jie Hwan Lee, Hoon Heo, Dong suk Jeon, Kyogu Lee

WEIGHTED MAGNITUDE-PHASE LOSS FOR SPEECH DEREVERBERATION ....................... 5794
  Jingshu Zhang, Mark D. Plumbley, Wenwu Wang

SPE-7: SPEAKER RECOGNITION 1: BENCHMARK EVALUATION

SPEAKER EMBEDDINGS FOR DIARIZATION OF BROADCAST DATA IN THE ALLIES
CHALLENGE ....................................................................................................................... 5799
  Anthony Larcher, Ambuj Mehrish, Marie Tahon, Sylvain Meignier, Jean Carrive, David
  Doukhan, Olivier Galibert, Nicholas Evans

ON THE DETECTION OF PITCH-SHIFTED VOICE: MACHINES AND HUMAN LISTENERS..... 5804
  David Looney, Nikolay D. Gaubitch

THE INS AND OUTS OF SPEAKER RECOGNITION: LESSONS FROM VOXSRC 2020 .......... 5809
  Yoohwan Kwon, Hee-Soo Heo, Bong-Jin Lee, Joon Son Chung

THE IDLAB VOXSRC-20 SUBMISSION: LARGE MARGIN FINE-TUNING AND QUALITY-
AWARE SCORE CALIBRATION IN DNN BASED SPEAKER VERIFICATION ....................... 5814
  Jenthe Thienpondt, Brecht Desplanques, Kris Demuynck
ANALYSIS OF THE BUT DIARIZATION SYSTEM FOR VOXCONVERSE CHALLENGE .............................. 5819
Federico Landini, Ondrej Glembek, Pavel Matejka, Johan Rohdin, Lukáš Burget, Mireita Diez, Anna Silnova

MICROSOFT SPEAKER DIARIZATION SYSTEM FOR THE VOXCELEB SPEAKER RECOGNITION CHALLENGE ............................... 5824
Xiong Xiao, Naoyuki Kanda, Zhuo Chen, Tianyan Zhou, Takuya Yoshiooka, Sanyuan Chen, Yong Zhao, Gang Liu, Yu Wu, Jian Wu, Shujie Liu, Jinyu Li, Yifan Gong

SPE-8: SPEAKER RECOGNITION 2: CHANNEL AND DOMAIN ROBUSTNESS

SQUEEZING VALUE OF CROSS-DOMAIN LABELS: A DECOUPLED SCORING APPROACH FOR SPEAKER VERIFICATION ................................................................. 5829
Lantian Li, Yang Zhang, Jiawen Kang, Thomas Fang Zheng, Dong Wang

SELF-SUPERVISED LEARNING BASED DOMAIN ADAPTATION FOR ROBUST SPEAKER VERIFICATION ................................................................................................. 5834
Zhengyang Chen, Shuai Wang, Yanmin Qian

META-LEARNING FOR CROSS-CHANNEL SPEAKER VERIFICATION ........................................................................................................... 5839
Hanyi Zhang, Longbiao Wang, Kong Aik Lee, Meng Liu, Jianwu Dang, Hui Chen

SYNAUG: SYNTHESIS-BASED DATA AUGMENTATION FOR TEXT-DEPENDENT SPEAKER VERIFICATION ................................................................................................. 5844
Chenpeng Du, Bing Han, Shuai Wang, Yanmin Qian, Kai Yu

UNIT SELECTION SYNTHESIS BASED DATA AUGMENTATION FOR FIXED PHRASE SPEAKER VERIFICATION ................................................................................................. 5849
Houjun Huang, Xu Xiang, Fei Zhao, Shuai Wang, Yanmin Qian

IMPROVING SPEAKER VERIFICATION IN REVERBERANT ENVIRONMENTS ................................................................. 5854
Xiao Chen, Stephen A. Zahorian

SPE-9: SPEECH RECOGNITION 3: TRANSFORMER MODELS 1

TRANSFORMER-TRANSDUCERS FOR CODE-SWITCHED SPEECH RECOGNITION ................................................................................................. 5859
Siddharth Dalmia, Yuzong Liu, Srikanth Ronanki, Katrin Kirchhoff

WAKE WORD DETECTION WITH STREAMING TRANSFORMERS ................................................................................................. 5864
Yiming Wang, Hang Lv, Daniel Povey, Lei Xie, Sanjeev Khudanpur

CAPTURING MULTI-RESOLUTION CONTEXT BY DILATED SELF-ATTENTION ................................................................................................. 5869
Niko Moritz, Takaaki Hori, Jonathan Le Roux

RECENT DEVELOPMENTS ON ESPNET TOOLKIT BOOSTED BY CONFORMER ................................................................................................. 5874
Pengcheng Guo, Florian Boyer, Xuankai Chang, Tomoki Hayashi, Yosuke Higuchi, Hirofumi Inaguma, Naoyuki Kamo, Chenda Li, Daniel Garcia-Romero, Jiatong Shi, Jing Shi, Shinji Watanabe, Kun Wei, Wangyou Zhang, Yuekai Zhang

HIERARCHICAL TRANSFORMER-BASED LARGE-CONTEXT END-TO-END ASR WITH LARGE-CONTEXT KNOWLEDGE DISTILLATION ................................................................................................. 5879
Ryo Masumura, Naoki Makishima, Mana Ihori, Akihiko Takashima, Tomohiro Tanaka, Shota Orihashi
END-TO-END MULTI-CHANNEL TRANSFORMER FOR SPEECH RECOGNITION ........................................... 5884
Feng-Ju Chang, Martin Radfar, Athanasios Mouchtaris, Brian King, Siegfried Kunzmann

SPE-10: SPEECH RECOGNITION 4: TRANSFORMER MODELS 2

CASS-NAT: CTC ALIGNMENT-BASED SINGLE STEP NON-AUTOREGRESSIVE TRANSFORMER FOR SPEECH RECOGNITION................................................................. 5889
Ruchao Fan, Wei Chu, Peng Chang, Jing Xiao

NON-AUTOREGRESSIVE TRANSFORMER ASR WITH CTC-ENHANCED DECODER INPUT ......................................................................................................................... 5894
Xingchen Song, Zhiyong Wu, Yiheng Huang, Chao Weng, Dan Su, Helen Meng

TRANSFORMER-BASED END-TO-END SPEECH RECOGNITION WITH LOCAL DENSE SYNTHESIZER ATTENTION ...................................................................................... 5899
Menglong Xu, Shengqiang Li, Xiao-Lei Zhang

DEVELOPING REAL-TIME STREAMING TRANSFORMER TRANSDUCER FOR SPEECH RECOGNITION ON LARGE-SCALE DATASET .................................................... 5904
Xie Chen, Yu Wu, Zhenghao Wang, Shujie Liu, Jinyu Li

HEAD-SYNCHRONOUS Decoding FOR TRANSFORMER-BASED STREAMING ASR.................. 5909
Mohan Li, Catalin Zorila, Rama Doddipatla

HISTORY UTTERANCE EMBEDDING TRANSFORMER LM FOR SPEECH RECOGNITION ........... 5914
Keqi Deng, Gaofeng Cheng, Haoran Miao, Pengyuan Zhang, Yonghong Yan

SPE-11: VOICE CONVERSION 1: NON-PARALLEL CONVERSION

MASKCYCLEGAN-VC: LEARNING NON-PARALLEL VOICE CONVERSION WITH FILLING IN FRAMES ................................................................................................. 5919
Takuhiro Kaneko, Hirokazu Kameoka, Kou Tanaka, Nobukatsu Hojo

NON-PARALLEL MANY-TO-MANY VOICE CONVERSION BY KNOWLEDGE TRANSFER FROM A TEXT-TO-SPEECH MODEL ......................................................... 5924
Xinyuan Yu, Brian Mak

NON-PARALLEL MANY-TO-MANY VOICE CONVERSION USING LOCAL LINGUISTIC TOKENS ............................................................ 5929
Chao Wang, Yibiao Yu

CRANK: AN OPEN-SOURCE SOFTWARE FOR NONPARALLEL VOICE CONVERSION BASED ON VECTOR-QUANTIZED VARIATIONAL AUTOENCODER........................................ 5934
Kazuhiro Kobayashi, Wen-Chin Huang, Yi-Chiaow Wu, Patrick Lumban Tobing, Tomoki Hayashi, Tomoki Toda

FRAGMENTVC: ANY-TO-ANY VOICE CONVERSION BY END-TO-END EXTRACTING AND FUSING FINE-GRAINED VOICE FRAGMENTS WITH ATTENTION ........................... 5939
Yist Y. Lin, Chung-Ming Chien, Jheng-Hao Lin, Hung-Yi Lee, Lin-Shan Lee

ANY-TO-ONE SEQUENCE-TO-SEQUENCE VOICE CONVERSION USING SELF-SUPERVISED DISCRETE SPEECH REPRESENTATIONS ....................................................... 5944
Wen-Chin Huang, Yi-Chiao Wu, Tomoki Hayashi
SPE-12: VOICE CONVERSION 2: LOW-RESOURCE & CROSS-LINGUAL CONVERSION

TOWARDS LOW-RESOURCE STARGAN VOICE CONVERSION USING WEIGHT ADAPTIVE INSTANCE NORMALIZATION ............................................................................................................................. 5949
Mingjie Chen, Yanpei Shi, Thomas Hain

AGAIN-VC: A ONE-SHOT VOICE CONVERSION USING ACTIVATION GUIDANCE AND ADAPTIVE INSTANCE NORMALIZATION ........................................................................................................ 5945
Yen-Hao Chen, Da-Yi Wu, Tsung-Han Wu, Hung-Yi Lee

ONE-SHOT VOICE CONVERSION BASED ON SPEAKER AWARE MODULE ............................................................................................................................. 5959
Ying Zhang, Hao Che, Jie Li, Chenxing Li, Xiaorui Wang, Zhongyuan Wang

ZERO-SHOT VOICE CONVERSION WITH ADJUSTED SPEAKER EMBEDDINGS AND SIMPLE ACOUSTIC FEATURES ................................................................................................................ 5964
Zhiyuan Tan, Jianguo Wei, Junhai Xu, Yuqing He, Wenhuan Lu

TOWARDS NATURAL AND CONTROLLABLE CROSS-LINGUAL VOICE CONVERSION BASED ON NEURAL TTS MODEL AND PHONETIC POSTERIORGRAM ........................................... 5969
Shengkui Zhao, Hao Wang, Trung Hieu Nguyen, Bin Ma

SPE-13: SPEECH RECOGNITION 5: NEW ALGORITHMS

META-LEARNING FOR IMPROVING RARE WORD RECOGNITION IN END-TO-END ASR ........... 5974
Florian Lux, Ngoc Thang Vu

A COMPARISON OF METHODS FOR OOV-WORD RECOGNITION ON A NEW PUBLIC DATASET...................................................................................................................................................... 5979
Rudolf A. Braun, Srikanth Madikeri, Petr Motlicek

CONVOLUTIONAL DROPOUT AND WORDPIECE AUGMENTATION FOR END-TO-END SPEECH RECOGNITION ............................................................................................................................. 5984
Hainan Xu, Yinghui Huang, Yun Zhu, Kartik Audhkhasi, Bhuvana Ramabhadran

PARTIALLY OVERLAPPED INFERENCE FOR LONG-FORM SPEECH RECOGNITION .............. 5989
Tae Gyoon Kang, Ho-Gyeong Kim, Min-Joong Lee, Jihyun Lee, Hoshik Lee

FOCUS ON THE PRESENT: A REGULARIZATION METHOD FOR THE ASR SOURCE-TARGET ATTENTION LAYER .................................................................................................................. 5994
Nanxin Chen, Piotr Zelasko, Jesús Villalba, Najim Dehak

BIFOCAL NEURAL ASR: EXPLOITING KEYWORD SPOTTING FOR INFERENCE OPTIMIZATION ............................................................................................................................. 5999
Jon Macoskey, Grant P. Strimel, Ariya Rastrow

SPE-14: SPEECH RECOGNITION 6: NEW ALGORITHMS FOR SPARSITY/EFFICIENCY

FASTEMIT: LOW-LATENCY STREAMING ASR WITH SEQUENCE-LEVEL EMISSION REGULARIZATION ............................................................................................................................. 6004
Jiahui Yu, Chung-Cheng Chiu, Bo Li, Shuo-Yiin Chang, Tara N. Sainath, Yanzhang He, Arun Narayanan, Wei Han, Anmol Gulati, Yonghui Wu, Ruoming Pang
SPARSIFICATION VIA COMPRESSED SENSING FOR AUTOMATIC SPEECH RECOGNITION............................................................................................................................................. 6009
    Kai Zhen, Hieu Duy Nguyen, Feng-Ju Chang, Athanasios Mouchtaris, Ariya Rastrow

DYNAMIC SPARSITY NEURAL NETWORKS FOR AUTOMATIC SPEECH RECOGNITION........ 6014
    Zhaofeng Wu, Ding Zhao, Qiao Liang, Jiahui Yu, Anmol Gulati, Ruoming Pang

AN ASYNCHRONOUS WFST-BASED DECODER FOR AUTOMATIC SPEECH RECOGNITION............................................................................................................................................. 6019
    Hang Lv, Zhehuai Chen, Hainan Xu, Daniel Povey, Lei Xie, Sanjeev Khudanpur

TINY TRANSDUCER: A HIGHLY-EFFICIENT SPEECH RECOGNITION MODEL ON EDGE DEVICES.......................................................................................................................................................................................... 6024
    Yuekai Zhang, Sining Sun, Long Ma

SPE-15: SPEECH SYNTHESIS 3: VOCODER

NOISE LEVEL LIMITED SUB-MODELING FOR DIFFUSION PROBABILISTIC VOCODERS........ 6029
    Takuma Okamoto, Tomoki Toda, Yoshinori Shiga, Hisashi Kawai

STYLEMELGAN: AN EFFICIENT HIGH-FIDELITY ADVERSARIAL VOCODER WITH TEMPORAL ADAPTIVE NORMALIZATION ............................................................................................................................................. 6034
    Ahmed Mustafa, Nicola Pia, Guillaume Fuchs

PARALLEL WAVEFORM SYNTHESIS BASED ON GENERATIVE ADVERSARIAL NETWORKS WITH VOICING-AWARE CONDITIONAL DISCRIMINATORS..................................................................................................................................... 6039
    Ryuichi Yamamoto, Eunwoo Song, Min-Jae Hwang, Jae-Min Kim

UNIVERSAL NEURAL VOCODING WITH PARALLEL WAVENET.......................................................................................................................................................... 6044
    Yunlong Jiao, Adam Gabrys, Georgi Tinchev, Bartosz Putrycz, Daniel Korzekwa, Viacheslav Klimkov

PERIODNET: A NON-AUTOREGRESSIVE WAVEFORM GENERATION MODEL WITH A STRUCTURE SEPARATING PERIODIC AND APERIODIC COMPONENTS..................................................................................................................................... 6049
    Yukiya Hono, Shinji Takaki, Kei Hashimoto, Keiichiro Oura, Yoshihiko Nankaku, Keiichi Tokuda

LVCNET: EFFICIENT CONDITION-DEPENDENT MODELING NETWORK FOR WAVEFORM GENERATION.......................................................................................................................................................... 6054
    Zhen Zeng, Jianzong Wang, Ning Cheng, Jing Xiao

SPE-16: SPEECH SYNTHESIS 4: FRONT-END

GRAPHSPEECH: SYNTAX-AWARE GRAPH ATTENTION NETWORK FOR NEURAL SPEECH SYNTHESIS...................................................................................................................................................... 6059
    Rui Liu, Berrak Sisman, Haizhou Li

SYNTACTIC REPRESENTATION LEARNING FOR NEURAL NETWORK BASED TTS WITH SYNTACTIC PARSE TREE TRAVERSAL..................................................................................................................................... 6064
    Changhe Song, Jingbei Li, Yixuan Zhou, Zhiyong Wu, Helen Meng

A CHAPTER-WISE UNDERSTANDING SYSTEM FOR TEXT-TO-SPEECH IN CHINESE NOVELS.......................................................................................................................................................... 6069
    Junjie Pan, Lin Wu, Xiang Yin, Pengfei Wu, Chengchang Xu, Zejun Ma
A UNIVERSAL BERT-BASED FRONT-END MODEL FOR MANDARIN TEXT-TO-SPEECH SYNTHESIS .................................................................................................................................................. 6074
Zilong Bai, Beibei Hu

IMPROVING PROSODY MODELLING WITH CROSS-UTTERANCE BERT EMBEDDINGS FOR END-TO-END SPEECH SYNTHESIS .................................................................................................................................................. 6079
Guanghui Xu, Wei Song, Zhengchen Zhang, Chao Zhang, Xiaodong He, Bowen Zhou

SPE-17: SPEECH ENHANCEMENT 3: TARGET SPEECH EXTRACTION

TIME-DOMAIN SPEECH EXTRACTION WITH SPATIAL INFORMATION AND MULTI SPEAKER CONDITIONING MECHANISM .................................................................................................................................................. 6084
Jisi Zhang, Catalin Zorila, Rama Doddipatla, Jon Barker

ADL-MVDR: ALL DEEP LEARNING MVDR BEAMFORMER FOR TARGET SPEECH SEPARATION .................................................................................................................................................. 6089
Zhuohuang Zhang, Yong Xu, Meng Yu, Shi-Xiong Zhang, Lianwu Chen, Dong Yu

MULTI-CHANNEL TARGET SPEECH EXTRACTION WITH CHANNEL DECORRELATION AND TARGET SPEAKER ADAPTATION .................................................................................................................................................. 6094
Jiangyu Han, Xinyuan Zhou, Yanhua Long, Yijie Li

SPEAKER ACTIVITY DRIVEN NEURAL SPEECH EXTRACTION .................................................................................................................................................. 6099
Marc Delcroix, Katerina Zmolikova, Tsubasa Ochiai, Keisuke Kinoshita, Tomohiro Nakatani

WASE: LEARNING WHEN TO ATTEND FOR SPEAKER EXTRACTION IN COCKTAIL PARTY ENVIRONMENTS .................................................................................................................................................. 6104
Yunzhe Hao, Jiaming Xu, Peng Zhang, Bo Xu

MULTI-STAGE SPEAKER EXTRACTION WITH UTTERANCE AND FRAME-LEVEL REFERENCE SIGNALS .................................................................................................................................................. 6109
Meng Ge, Chenglin Xu, Longbiao Wang, Eng Siong Chng, Jianwu Dang, Haizhou Li

SPE-18: SPEECH ENHANCEMENT 4: MULTI-CHANNEL PROCESSING

NEURAL NETWORK-BASED VIRTUAL MICROPHONE ESTIMATOR .................................................................................................................................................. 6114
Tsubasa Ochiai, Marc Delcroix, Tomohiro Nakatani, Rintaro Ikeshita, Keisuke Kinoshita, Shoko Araki

JOINT MAXIMUM LIKELIHOOD ESTIMATION OF POWER SPECTRAL DENSITIES AND RELATIVE ACOUSTIC TRANSFER FUNCTIONS FOR ACOUSTIC BEAMFORMING .................................................................................................................................................. 6119
Poul Hoang, Zheng-Hua Tan, Jan Mark De Haan, Jesper Jensen

CUE-PRESERVING MMSE FILTER WITH BAYESIAN SNR MARGINALIZATION FOR BINAURAL SPEECH ENHANCEMENT .................................................................................................................................................. 6124
Stefan Thaleisier, Gerald Enzner

BLIND AND NEURAL NETWORK-GUIDED CONVOLUTIONAL BEAMFORMER FOR JOINT DENOISING, DEREVERBERATION, AND SOURCE SEPARATION .................................................................................................................................................. 6129
Tomohiro Nakatani, Rintaro Ikeshita, Keisuke Kinoshita, Hiroshi Sawada, Shoko Araki

REAL-TIME SPEECH ENHANCEMENT FOR MOBILE COMMUNICATION BASED ON DUAL-CHANNEL COMPLEX SPECTRAL MAPPING .................................................................................................................................................. 6134
Ke Tan, Xueliang Zhang, Deliang Wang
DON’T SHOOT BUTTERFLY WITH RIFLES: MULTI-CHANNEL CONTINUOUS SPEECH SEPARATION WITH EARLY EXIT TRANSFORMER ................................................................. 6139
   Sanyuan Chen, Yu Wu, Zhuo Chen, Takuya Yoshioka, Shujie Liu, Jinyu Li, Xiangzhan Yu

SPE-19: SPEAKER RECOGNITION 3: ATTENTION AND ADVERSARIAL

DOUBLE MULTI-HEAD ATTENTION FOR SPEAKER VERIFICATION ........................................ 6144
   Miquel India, Pooyan Safari, Javier Hernando

GRAPH ATTENTION NETWORKS FOR SPEAKER VERIFICATION........................................... 6149
   Jee-Weon Jung, Hee-Soo Heo, Ha-Jin Yu, Joon Son Chung

MEMORY LAYERS WITH MULTI-HEAD ATTENTION MECHANISMS FOR TEXT-DEPENDENT SPEAKER VERIFICATION ................................................................. 6154
   Victoria Mingote, Antonio Miguel, Alfonso Ortega, Eduardo Lleida

FOOLHD: FOOLING SPEAKER IDENTIFICATION BY HIGHLY IMPERCEPTIBLE ADVERSARIAL DISTURBANCES................................................................. 6159
   Ali Shahin Shamsabadi, Francisco Sepúlveda Teixeira, Alberto Abad, Bhiksha Raj, Andrea Cavallaro, Isabel Trancoso

ADVERSARIAL DEFENSE FOR DEEP SPEAKER RECOGNITION USING HYBRID ADVERSARIAL TRAINING ................................................................. 6164
   Monisankha Pal, Arindam Jati, Raghavree Peri, Chin-Cheng Hsu, Wael Abdalmageed, Shrikanth Narayanan

DEAAN: DISENTANGLLED EMBEDDING AND ADVERSARIAL ADAPTATION NETWORK FOR ROBUST SPEAKER REPRESENTATION LEARNING ........................................... 6169
   Mufan Sang, Wei Xia, John H. L. Hansen

SPE-20: SPEAKER RECOGNITION 4: APPLICATIONS

PLAYING A PART: SPEAKER VERIFICATION AT THE MOVIES............................................... 6174
   Andrew Brown, Jaesung Huh, Arsha Nagrani, Joon Son Chung, Andrew Zisserman

SMALL FOOTPRINT TEXT-INDEPENDENT SPEAKER VERIFICATION FOR EMBEDDED SYSTEMS ............................................................................ 6179
   Julien Balian, Raffaele Tavarone, Mathieu Pounyrol, Alice Coucke

ASV-SUBTOOLS: OPEN SOURCE TOOLKET FOR AUTOMATIC SPEAKER VERIFICATION......... 6184
   Fuchuan Tong, Miao Zhao, Jianfeng Zhou, Hao Lu, Zheng Li, Lin Li, Qingyang Hong

DEEPTALK: VOCAL STYLE ENCODING FOR SPEAKER RECOGNITION AND SPEECH SYNTHESIS ................................................................. 6189
   Anurag Chowdhury, Arun Ross, Prabu David

A MULTI-VIEW APPROACH TO AUDIO-VISUAL SPEAKER VERIFICATION .............................. 6194
   Leda Sari, Kritika Singh, Jiatao Zhou, Lorenzo Torresani, Nayan Singhal, Yatharth Saraf

SPE-21: SPEECH RECOGNITION 7: TRAINING METHODS FOR END-TO-END MODELING

TOP-DOWN ATTENTION IN END-TO-END SPOKEN LANGUAGE UNDERSTANDING .................. 6199
   Yixin Chen, Weiyi Lu, Alejandro Mottini, Li Erran Li, Jasha Droppo, Zheng Du, Belinda Zeng
FINE-TUNING OF PRE-TRAINED END-TO-END SPEECH RECOGNITION WITH GENERATIVE ADVERSARIAL NETWORKS
Md. Akmal Haidar, Mehdi Rezagholizadeh

A GENERAL MULTI-TASK LEARNING FRAMEWORK TO LEVERAGE TEXT DATA FOR SPEECH TO TEXT TASKS
Yun Tang, Juan Pino, Changhan Wang, Xutai Ma, Dmitriy Genzel

GAUSSIAN KERNELIZED SELF-ATTENTION FOR LONG SEQUENCE DATA AND ITS APPLICATION TO CTC-BASED SPEECH RECOGNITION
Yosuke Kashiwagi, Emiru Tsunoo, Shinji Watanabe

LATTICE-FREE MMI ADAPTATION OF SELF-SUPERVISED PRETRAINED ACOUSTIC MODELS
Apoorv Vyas, Srikanth Madikeri, Hervé Bourlard

INTERMEDIATE LOSS REGULARIZATION FOR CTC-BASED SPEECH RECOGNITION
Jaesong Lee, Shinji Watanabe

**SPE-22: SPEECH RECOGNITION 8: MULTILINGUAL SPEECH RECOGNITION**

CODE-SWITCH SPEECH RESCORING WITH MONOLINGUAL DATA
Guoyu Liu, Lixin Cao

MIXTURE OF INFORMED EXPERTS FOR MULTILINGUAL SPEECH RECOGNITION
Neeraj Gaur, Brian Farris, Parisa Haghani, Isabel Leal, Pedro J. Moreno, Manasa Prasad, Bhuvana Ramabhadran, Yun Zhu

REDUCING SPELLING INCONSISTENCIES IN CODE-SWITCHING ASR USING CONTEXTUALIZED CTC LOSS
Burin Naowarat, Thananchai Kongthaworn, Korrawe Karunratanakul, Sheng Hui Wu, Ekapol Chuangsuvanich

MULTI-DIALECT SPEECH RECOGNITION IN ENGLISH USING ATTENTION ON ENSEMBLE OF EXPERTS
Amit Das, Kshitiz Kumar, Jian Wu

DECOUPLING PRONUNCIATION AND LANGUAGE FOR END-TO-END CODE-SWITCHING AUTOMATIC SPEECH RECOGNITION
Shuai Zhang, Jiangyan Yi, Zhengkun Tian, Ye Bai, Jianhua Tao, Zhengqi Wen

AISPEECH-SITU ACCENT IDENTIFICATION SYSTEM FOR THE ACCENTED ENGLISH SPEECH RECOGNITION CHALLENGE
Houjun Huang, Xu Xiang, Yexin Yang, Rao Ma, Yanmin Qian

**SPE-23: SPEECH EMOTION 1: SPEECH EMOTION RECOGNITION**

META-LEARNING FOR LOW-RESOURCE SPEECH EMOTION RECOGNITION
Suransh Chopra, Puneet Mathur, Ramit Sawhney, Rajiv Ratn Shah

PROGRESSIVE CO-TEACHING FOR AMBIGUOUS SPEECH EMOTION RECOGNITION
Yifei Yin, Yu Gu, Longshan Yao, Ying Zhou, Xuefeng Liang, He Zhang
EMOTION RECOGNITION BY FUSING TIME SYNCHRONOUS AND TIME ASYNCHRONOUS REPRESENTATIONS ................................................................. 6269
   Wen Wu, Chao Zhang, Philip C. Woodland

SPEECH EMOTION RECOGNITION BASED ON LISTENER ADAPTIVE MODELS ........................................... 6274
   Atsushi Ando, Ryo Masumura, Hiroshi Sato, Takaumi Moriya, Takanaori Ashihara, Yusuke Ijima, Tomoki Toda

SPEECH EMOTION RECOGNITION USING SEMANTIC INFORMATION ........................................... 6279
   Panagiotis Tzirakis, Anh Nguyen, Stefanos Zafeiriou, Björn W. Schuller

COMPACT GRAPH ARCHITECTURE FOR SPEECH EMOTION RECOGNITION ........................................... 6284
   Amir Shirian, Tanaya Guha

SPE-24: SPEECH EMOTION 2: NEURAL NETWORKS FOR SPEECH EMOTION RECOGNITION
A NOVEL END-TO-END SPEECH EMOTION RECOGNITION NETWORK WITH STACKED TRANSFORMER LAYERS ................................................................. 6289
   Xianfeng Wang, Min Wang, Wenbo Qi, Wanqi Su, Xiangqian Wang, Huan Zhou

A NOVEL ATTENTION-BASED GATED RECURRENT UNIT AND ITS EFFICACY IN SPEECH EMOTION RECOGNITION ................................................................. 6294
   Srvidya Tirunellai Rajamani, Kumar T. Rajamani, Adria Mallol-Ragolta, Shuo Liu, Björn Schuller

MAEC: MULTI-INSTANCE LEARNING WITH AN ADVERSARIAL AUTO-ENCODER-BASED CLASSIFIER FOR SPEECH EMOTION RECOGNITION ................................................................. 6299
   Changzeng Fu, Chaoran Liu, Carlos Toshinori Ishi, Hiroshi Ishiguro

REPRESENTATION LEARNING WITH SPECTRO-TEMPORAL-CHANNEL ATTENTION FOR SPEECH EMOTION RECOGNITION ................................................................. 6304
   Lili Guo, Longbiao Wang, Chenglin Xu, Jianwu Dang, Eng Siong Chng, Haizhou Li

SPEECH EMOTION RECOGNITION USING QUATERNION CONVOLUTIONAL NEURAL NETWORKS ................................................................. 6309
   Aneesh Muppidi, Martin Radfar

DOMAIN-ADVERSARIAL AUTOENCODER WITH ATTENTION BASED FEATURE LEVEL FUSION FOR SPEECH EMOTION RECOGNITION ................................................................. 6314
   Yuan Gao, Jiaxing Liu, Longbiao Wang, Jianwu Dang

SPE-25: SPEECH EMOTION 3: EMOTION RECOGNITION - REPRESENTATIONS, DATA AUGMENTATION
SPEECH EMOTION RECOGNITION WITH MULTISCALE AREA ATTENTION AND DATA AUGMENTATION ................................................................. 6319
   Mingke Xu, Fan Zhang, Xiaodong Cui, Wei Zhang

COPYPASTE: AN AUGMENTATION METHOD FOR SPEECH EMOTION RECOGNITION ...................... 6324
   Raghavendra Pappagari, Jesús Villalba, Piotr Zelasko, Laureano Moro-Velazquez, Najim Dehak
CONTRASTIVE UNSUPERVISED LEARNING FOR SPEECH EMOTION RECOGNITION

Mao Li, Bo Yang, Joshua Levy, Andreas Stolcke, Viktor Rozgic, Spyros Matsoukas, Constantinos Papayiannis, Daniel Bone, Chao Wang

HIERARCHICAL NETWORK BASED ON THE FUSION OF STATIC AND DYNAMIC FEATURES FOR SPEECH EMOTION RECOGNITION

Qi Cao, Mixiao Hou, Bingzhi Chen, Zheng Zhang, Guangming Lu

MULTIMODAL EMOTION RECOGNITION WITH CAPSULE GRAPH CONVOLUTIONAL BASED REPRESENTATION FUSION

Jiaxing Liu, Sen Chen, Longbiao Wang, Zhilei Liu, Yahui Fu, Lili Guo, Jianwu Dang

DISENTANGLEMENT FOR AUDIO-VISUAL EMOTION RECOGNITION USING MULTITASK SETUP

Raghuveer Peri, Srinivas Parthasarathy, Charles Bradshaw, Shiva Sundaram

SPE-26: SPEAKER VERIFICATION SPOOFING AND COUNTERMEASURES

DATA AUGMENTATION WITH SIGNAL COMPANDING FOR DETECTION OF LOGICAL ACCESS ATTACKS

Rohan Kumar Das, Jichen Yang, Haizhou Li

REPLAY AND SYNTHETIC SPEECH DETECTION WITH RES2NET ARCHITECTURE

Xu Li, Na Li, Chao Weng, Xunying Liu, Dan Su, Dong Yu, Helen Meng

A CAPSULE NETWORK BASED APPROACH FOR DETECTION OF AUDIO SPOOFING ATTACKS

Anwei Luo, Enlei Li, Yongliang Liu, Xiangui Kang, Z. Jane Wang

CROSS-TEAGER ENERGY CEPSTRAL COEFFICIENTS FOR REPLAY SPOOF DETECTION ON VOICE ASSISTANTS

Rajul Acharya, Harsh Kotta, Ankur T. Patil, Hemant A. Patil

END-TO-END ANTI-SPOOFING WITH RAWNET2

Hemlata Tak, Jose Patino, Massimiliano Todisco, Andreas Nautsch, Nicholas Evans, Anthony Larcher

REPLAY-ATTACK DETECTION USING FEATURES WITH ADAPTIVE SPECTRO-TEMPORAL RESOLUTION

Meng Liu, Longbiao Wang, Kong Aik Lee, Xuanda Chen, Jianwu Dang

SPE-27: SPEECH RECOGNITION 9: CONFIDENCE MEASURES

IMPROVING IDENTIFICATION OF SYSTEM-DIRECTED SPEECH UTTERANCES BY DEEP LEARNING OF ASR-BASED WORD EMBEDDINGS AND CONFIDENCE METRICS

Vilayphone Vilaysouk, Amr Nour-Eldin, Dermot Connolly

BLSTM-BASED CONFIDENCE ESTIMATION FOR END-TO-END SPEECH RECOGNITION

Atsunori Ogawa, Naohiro Tawara, Takatomo Kano, Marc Delcroix

CONFIDENCE ESTIMATION FOR ATTENTION-BASED SEQUENCE-TO-SEQUENCE MODELS FOR SPEECH RECOGNITION

Qiujia Li, David Qiu, Yu Zhang, Bo Li, Yanzhang He, Philip C. Woodland, Liangliang Cao, Trevor Strohman
LEARNING WORD-LEVEL CONFIDENCE FOR SUBWORD END-TO-END ASR
David Qiu, Qiujia Li, Yanzhang He, Yu Zhang, Bo Li, Liangliang Cao, Rohit Prabhavalkar,
Deepti Bhatia, Wei Li, Ke Hu, Tara N. Sainath, Ian McGraw

NEURAL UTERANCE CONFIDENCE MEASURE FOR RNN-TRANSUDERS AND TWO
PASS MODELS
Ashutosh Gupta, Ankur Kumar, Dhananjaya Gowda, Kwangyoun Kim, Sachin Singh,
Shatrughan Singh, Chanwoo Kim

DETECTING ADVERSARIAL ATTACKS ON AUDIOVISUAL SPEECH RECOGNITION
Pingchuan Ma, Stavros Petridis, Maja Pantic

SPE-28: SPEECH RECOGNITION 10: ROBUSTNESS TO HUMAN SPEECH VARIABILITY

REDAT: ACCENT-INVARIANT REPRESENTATION FOR END-TO-END ASR BY DOMAIN
ADVERSARIAL TRAINING WITH RELABELING
Hu Hu, Xuesong Yang, Zeynab Raeesy, Jinxi Guo, Gokce Keskin, Harish Arskere, Ariya
Rastrow, Andreas Stolcke, Roland Maas

AISPEECH-SJTU ASR SYSTEM FOR THE ACCENTED ENGLISH SPEECH RECOGNITION
CHALLENGE
Tian Tan, Yizhou Lu, Rao Ma, Sen Zhu, Jiaqi Guo, Yanmin Qian

END-TO-END MULTI-ACCENT SPEECH RECOGNITION WITH UNSUPERVISED ACCENT
MODELLING
Song Li, Beibei Ouyang, Dexin Liao, Shipeng Xia, Lin Li, Qingyang Hong

A COMPARATIVE STUDY OF ACOUSTIC AND LINGUISTIC FEATURES
CLASSIFICATION FOR ALZHEIMER'S DISEASE DETECTION
Jinchao Li, Jianwei Yu, Zi Ye, Simon Wong, Manwai Mak, Brian Mak, Xunying Liu, Helen
Meng

SYNTHESIS OF NEW WORDS FOR IMPROVED DYSAURALIC SPEECH RECOGNITION
ON AN EXPANDED VOCABULARY
John Harvill, Dias Issa, Mark Hasegawa-Johnson, Changdong Yoo

DEVELOPMENT OF THE CUHK ELDERLY SPEECH RECOGNITION SYSTEM FOR
NEUROCOGNITIVE DISORDER DETECTION USING THE DEMENTIABANK CORPUS
Zi Ye, Shoukang Hu, Jinchao Li, Xurong Xie, Mengzhe Geng, Jianwei Yu, Junhao Xu, Boyang
Xue, Shansong Liu, Xunying Liu, Helen Meng

SPE-29: SPEECH PROCESSING 1: PRODUCTION

PORTABLE PHOTOGLOTTOGRAPHY FOR MONITORING VOCAL FOLD VIBRATIONS IN
SPEECH PRODUCTION
Yujie Chi, Kiyoshi Honda, Jianguo Wei

IMPROVING ULTRASOUND TONGUE CONTOUR EXTRACTION USING U-NET AND
SHAPE CONSISTENCY-BASED REGULARIZER
Ming Feng, Yin Wang, Kele Xu, Huaimin Wang, Bo Ding

IMPACT OF SPEAKING RATE ON THE SOURCE FILTER INTERACTION IN SPEECH: A
STUDY
Tilak Purohit, MV Achuth Rao, Prasanta Kumar Ghosh
A TWO-STAGE DEEP MODELING APPROACH TO ARTICULATORY INVERSION
Abdolreza Sabzi Shahrebabaki, Negar Olfati, Ali Shariq Imran, Magne Hallstein Johnsen,
Sabato Marco Siniscalchi, Torbjørn Svendsen

ACOUSTIC-TO-ARTICULATORY INVERSION FOR DYSARTHRIC SPEECH BY USING CROSS-CORPUS ACOUSTIC-ARTICULATORY DATA
Sarthak Kumar Maharana, Aravind Illa, Renuka Mannem, Yamini Belur, Preetie Shetty,
Veeramani Preethish Kumar, Seena Vengalil, Kiran Polavarapu, Nalini Atchayaram,
Prasanta Kumar Ghosh

SPEAKING RATE AND TONAL REALIZATION IN MANDARIN CHINESE: WHAT CAN WE LEARN FROM LARGE SPEECH CORPORA?
Jiahong Yuan, Kenneth Church

SPE-30: SPEECH PROCESSING 2: GENERAL TOPICS
HUMANACGAN: CONDITIONAL GENERATIVE ADVERSARIAL NETWORK WITH HUMAN-BASED AUXILIARY CLASSIFIER AND ITS EVALUATION IN PHONEME PERCEPTION
Yota Ueda, Kazuki Fujii, Yuki Saito, Shinnosuke Takamichi, Yukino Baba, Hiroshi Saruwatari

IMPROVING AUDIO ANOMALIES RECOGNITION USING TEMPORAL CONVOLUTIONAL ATTENTION NETWORKS
Qiang Huang, Thomas Hain

GENERATIVE SPEECH CODING WITH PREDICTIVE VARIANCE REGULARIZATION
W. Bastiaan Kleijn, Andrew Storus, Michael Chinen, Tom Denton, Felicia S. C. Lim,
Alejandro Luebs, Jan Skoglund, Hengchin Yeh

HOW TO MAKE TEXT-TO-SPEECH SYSTEM PRONOUNCE "VOLDEMORT": AN EXPERIMENTAL APPROACH OF FOREIGN WORD PHONEMIZATION IN VIETNAMESE
Dang-Khoa Mac, Van-Huy Nguyen, Dinh-Nghi Nguyen, Kim-Anh Nguyen

HOW SIMILAR OR DIFFERENT IS RAKUGO SPEECH SYNTHESIZER TO PROFESSIONAL PERFORMERS?
Shuhei Kato, Yusuke Yasuda, Xin Wang, Erica Cooper, Junichi Yamagishi

DNSMOS: A NON-INTRUSIVE PERCEPTUAL OBJECTIVE SPEECH QUALITY METRIC TO EVALUATE NOISE SUPPRESSORS
Chandan K A Reddy, Vishak Gopal, Ross Cutler

A CAUSAL DEEP LEARNING FRAMEWORK FOR CLASSIFYING PHONEMES IN COCHLEAR IMPLANTS
Kevin Chu, Leslie Collins, Boyla Mainsah

SPE-31: SPEECH RECOGNITION 11: NOVEL APPROACHES
MINIMUM BAYES RISK TRAINING FOR END-TO-END SPEAKER-ATTRIBUTED ASR
Naoyuki Kanda, Zhong Meng, Liang Lu, Yashesh Gaur, Xiaofei Wang, Zhou Chen, Takuya Yoshioka

MUTUALLY-CONSTRAINED MONOTONIC MULTIHEAD ATTENTION FOR ONLINE ASR
Jaeyun Song, Hajin Shim, Eunho Yang
THE USE OF VOICE SOURCE FEATURES FOR SUNG SPEECH RECOGNITION ........................................ 6513
Gerardo Roa Dabike, Jon Barker

A PARALLELIZABLE LATTICE RESCORING STRATEGY WITH NEURAL LANGUAGE MODELS ..................................................................................................................... 6518
Ke Li, Daniel Povey, Sanjeev Khudanpur

DECENTRALIZING FEATURE EXTRACTION WITH QUANTUM CONVOLUTIONAL NEURAL NETWORK FOR AUTOMATIC SPEECH RECOGNITION ...................................................... 6523
Chao-Han Huck Yang, Jun Qi, Samuel Yen-Chi Chen, Pin-Yu Chen, Sabato Marco Siniscalchi, Xiaoli Ma, Chin-Hui Lee

CIF-BASED COLLABORATIVE DECODING FOR END-TO-END CONTEXTUAL SPEECH RECOGNITION ................................................................. 6528
Minglun Han, Linkao Dong, Shiyu Zhou, Bo Xu

SPE-32: SPEECH RECOGNITION 12: SELF-SUPERVISED, SEMI-SUPERVISED, UNSUPERVISED TRAINING

HUBERT: HOW MUCH CAN A BAD TEACHER BENEFIT ASR PRE-TRAINING? ........................................ 6533
Wei-Ning Hsu, Yao-Hung Hubert Tsai, Benjamin Bolte, Ruslan Salakhutdinov, Abdelrahman Mohamed

A FURTHER STUDY OF UNSUPERVISED PRETRAINING FOR TRANSFORMER BASED SPEECH RECOGNITION ............................................................................................................................. 6538
Dongwei Jiang, Wubo Li, Ruixiong Zhang, Miao Cao, Ne Luo, Yang Han, Wei Zou, Kun Han, Xiangang Li

PRE-TRAINING TRANSFORMER DECODER FOR END-TO-END ASR MODEL WITH UNPAIRED TEXT DATA ............................................................................................................................. 6543
Changfeng Gao, Gaofeng Cheng, Runyan Yang, Han Zhu, Pengyuan Zhang, Yonghong Yan

SEMI-SUPERVISED SPEECH RECOGNITION VIA GRAPH-BASED TEMPORAL CLASSIFICATION ............................................................................................................................. 6548
Niko Moritz, Takaaki Hori, Jonathan Le Roux

UNSUPERVISED DOMAIN ADAPTATION FOR SPEECH RECOGNITION VIA UNCERTAINTY DRIVEN SELF-TRAINING ............................................................................................................................. 6553
Sameer Khurana, Niko Moritz, Takaaki Hori, Jonathan Le Roux

IMPROVING STREAMING AUTOMATIC SPEECH RECOGNITION WITH NON-STREAMING MODEL DISTILLATION ON UNSUPERVISED DATA ........................................................................ 6558
Thibault Doutre, Wei Han, Min Ma, Zhiyun Lu, Chung-Cheng Chiu, Ruoming Pang, Arun Narayanan, Ananya Misra, Yu Zhang, Liangliang Cao

SPE-33: SPEECH SYNTHESIS 5: PROSODY & STYLE

SPEECH BERT EMBEDDING FOR IMPROVING PROSODY IN NEURAL TTS ........................................ 6563
Liping Chen, Yan Deng, Xi Wang, Frank K. Soong, Lei He

BI-LEVEL STYLE AND PROSODY DECOUPLING MODELING FOR PERSONALIZED END-TO-END SPEECH SYNTHESIS ............................................................................................................................. 6568
Ruibo Fu, Jianhua Tao, Zhengqi Wen, Jiangyan Yi, Tao Wang, Chunyu Qiang
PROSODIC REPRESENTATION LEARNING AND CONTEXTUAL SAMPLING FOR NEURAL TEXT-TO-SPEECH ........................................................... 6573
Sri Karlapati, Ammar Abbas, Zack Hodari, Alexis Moinet, Arnaud Joly, Penny Karanasou, Thomas Drugman

CAMP: A TWO-STAGE APPROACH TO MODELLING PROSODY IN CONTEXT ............................................................... 6578
Zack Hodari, Alexis Moinet, Sri Karlapati, Jaime Lorenzo-Trueba, Thomas Merritt, Arnaud Joly, Ammar Abbas, Penny Karanasou, Thomas Drugman

UNSUPERVISED LEARNING FOR MULTI-STYLE SPEECH SYNTHESIS WITH LIMITED DATA ............................................................... 6583
Shuang Liang, Chenfeng Miao, Minchuan Chen, Jun Ma, Shaojun Wang, Jing Xiao

FASTPITCH: PARALLEL TEXT-TO-SPEECH WITH PITCH PREDICTION ............................................................... 6588
Adrian Lancucki

SPE-34: SPEECH SYNTHESIS 6: DATA AUGMENTATION & ADAPTATION

LOW-RESOURCE EXPRESSIVE TEXT-TO-SPEECH USING DATA AUGMENTATION .................................................. 6593
Goeric Huybrechts, Thomas Merritt, Giulia Comini, Bartek Perz, Raahil Shah, Jaime Lorenzo-Trueba

TTS-BY-TTS: TTS-DRIVEN DATA AUGMENTATION FOR FAST AND HIGH-QUALITY SPEECH SYNTHESIS .......................... 6598
Min-Jae Hwang, Ryuichi Yamamoto, Eunwoo Song, Jae-Min Kim

A NEURAL TEXT-TO-SPEECH MODEL UTILIZING BROADCAST DATA MIXED WITH BACKGROUND MUSIC ............................................ 6603
Hanbin Bae, Jae-Sung Bae, Young-Sun Joo, Young-Ik Kim, Hoon-Young Cho

DISENTANGLLED SPEAKER AND LANGUAGE REPRESENTATIONS USING MUTUAL INFORMATION MINIMIZATION AND DOMAIN ADAPTATION FOR CROSS-LINGUAL TTS .................................................................................. 6608
Detai Xin, Tatsuya Komatsu, Shinnosuke Takamichi, Hiroshi Saruwatari

ADASPEECH 2: ADAPTIVE TEXT TO SPEECH WITH UNTRANSCRIBED DATA ............................................................... 6613
Yuzi Yan, Xu Tan, Bohan Li, Tao Qin, Sheng Zhao, Yuan Shen, Tie-Yan Liu

INVESTIGATION OF FAST AND EFFICIENT METHODS FOR MULTI-SPEAKER MODELING AND SPEAKER ADAPTATION .......................... 6618
Yibin Zheng, Xinhui Li, Li Lu

SPE-35: SPEECH ENHANCEMENT 5: DNS CHALLENGE TASK

ICASSP 2021 DEEP NOISE SUPPRESSION CHALLENGE ........................................................................ 6623

ICASSP 2021 DEEP NOISE SUPPRESSION CHALLENGE: DECOUPLING MAGNITUDE AND PHASE OPTIMIZATION WITH A TWO-STAGE DEEP NETWORK .................................................. 6628
Andong Li, Wenzhe Liu, Xiaoxue Luo, Chengshi Zheng, Xiaodong Li

FULLSUBNET: A FULL-BAND AND SUB-BAND FUSION MODEL FOR REAL-TIME SINGLE-CHANNEL SPEECH ENHANCEMENT ........................................ 6633
Xiang Hao, Xiangdong Su, Radu Horaud, Xiaofei Li
DENSELY CONNECTED MULTI-STAGE MODEL WITH CHANNEL WISE SUBBAND FEATURE FOR REAL-TIME SPEECH ENHANCEMENT .......................................................... 6638
    Jingdong Li, Dawei Luo, Yun Liu, Yuanyuan Zhu, Zhaoxia Li, Guohui Cui, Wenqi Tang, Wei Chen

A MODULATION-DOMAIN LOSS FOR NEURAL-NETWORK-BASED REAL-TIME SPEECH ENHANCEMENT ............................................................. 6643
    Tyler Vuong, Yangyang Xia, Richard M. Stern

MONAURAL SPEECH ENHANCEMENT WITH COMPLEX CONVOLUTIONAL BLOCK ATTENTION MODULE AND JOINT TIME FREQUENCY LOSSES ...................................... 6648
    Shengkui Zhao, Trung Hieu Nguyen, Bin Ma

SPE-36: SPEECH ENHANCEMENT 6: MULTI-MODAL PROCESSING

AUDIO-VISUAL SPEECH INPAINTING WITH DEEP LEARNING .......................................................... 6653
    Giovanni Morrone, Daniel Michelsanti, Zheng-Hua Tan, Jesper Jensen

VSET: A MULTIMODAL TRANSFORMER FOR VISUAL SPEECH ENHANCEMENT .................................. 6658
    Karthik Ramesh, Chao Xing, Wupeng Wang, Dong Wang, Xiao Chen

SWITCHING VARIATIONAL AUTO-ENCODERS FOR NOISE-AGNOSTIC AUDIO-VISUAL SPEECH ENHANCEMENT ............................................................ 6663
    Mostafa Sadeghi, Xavier Alameda-Pineda

AUDIO-VISUAL SPEECH ENHANCEMENT METHOD CONDITIONED IN THE LIP MOTION AND SPEAKER-DISCRIMINATIVE EMBEDDINGS .............................................. 6668
    Koichiro Ito, Masaaki Yamamoto, Kenji Nagamatsu

AUDIO-VISUAL SPEECH SEPARATION USING CROSS-MODAL CORRESPONDENCE LOSS .................. 6673
    Naoki Makishima, Mana Ihori, Akihiko Takashima, Tomohiro Tanaka, Shota Orihashi, Ryo Masumura

MUSE: MULTI-MODAL TARGET SPEAKER EXTRACTION WITH VISUAL CUES .................................. 6678
    Zexu Pan, Ruijie Tao, Chenglin Xu, Haizhou Li

SPE-37: SPEAKER RECOGNITION 5: NEURAL EMBEDDING

AN EFFECTIVE DEEP EMBEDDING LEARNING METHOD BASED ON DENSE-RESIDUAL NETWORKS FOR SPEAKER VERIFICATION ........................................ 6683
    Ying Liu, Yan Song, Ian McLoughlin, Lin Liu, Li-Rong Dai

TIME-DOMAIN SPEAKER VERIFICATION USING TEMPORAL CONVOLUTIONAL NETWORKS .................. 6688
    Sangwook Han, Jaeuk Byun, Jong Won Shin

TOWARDS ROBUST SPEAKER VERIFICATION WITH TARGET SPEAKER ENHANCEMENT .................. 6693
    Chunlei Zhang, Meng Yu, Chao Weng, Dong Yu

A JOINT TRAINING FRAMEWORK OF MULTI-LOOK SEPARATOR AND SPEAKER EMBEDDING EXTRACTOR FOR OVERLAPPED SPEECH ........................................ 6698
    Naijun Zheng, Na Li, Bo Wu, Meng Yu, Jianwei Yu, Chao Weng, Dan Su, Xunying Liu, Helen Meng
CAM: CONTEXT-AWARE MASKING FOR ROBUST SPEAKER VERIFICATION ........................................ 6703
Ya-Qi Yu, Siqi Zheng, Hongbin Suo, Yun Lei, Wu-Jun Li

SHORT-TIME SPECTRAL AGGREGATION FOR SPEAKER EMBEDDING ........................................ 6708
Youzhi Tu, Man-Wai Mak

**SPE-38: SPEAKER RECOGNITION 6: SELF-SUPERVISED AND UNSUPERVISED LEARNING**

CONTRASTIVE SELF-SUPERVISED LEARNING FOR TEXT-INDEPENDENT SPEAKER VERIFICATION .................................................................................................................. 6713
Haoran Zhang, Yuexian Zou, Helin Wang

ADVERSARIAL DEFENSE FOR AUTOMATIC SPEAKER VERIFICATION BY CASCADED SELF-SUPERVISED LEARNING MODELS .................................................................................................. 6718
Haibin Wu, Xu Li, Andy T. Liu, Zhiyong Wu, Helen Meng, Hung-Yi Lee

SELF-SUPERVISED TEXT-DEPENDENT SPEAKER VERIFICATION USING PROTOTYPICAL MOMENTUM CONTRASTIVE LEARNING ................................................................. 6723
Wei Xia, Chunlei Zhang, Chao Weng, Meng Yu, Dong Yu

AN ITERATIVE FRAMEWORK FOR SELF-SUPERVISED DEEP SPEAKER REPRESENTATION LEARNING ................................................................................................................... 6728
Danwei Cai, Weiqing Wang, Ming Li

IMPROVING RECONSTRUCTION LOSS BASED SPEAKER EMBEDDING IN UNSUPERVISED AND SEMI-SUPERVISED SCENARIOS ............................................................................. 6733
Jaejin Cho, Piotr Zelasko, Jesús Villalba, Najim Dehak

**SPE-39: SPEECH RECOGNITION 13: ACOUSTIC MODELING 1**

SPEECH ACOUSTIC MODELLING FROM RAW PHASE SPECTRUM ............................................. 6738
Erfan Loweimi, Zoran Cvetkovic, Peter Bell, Steve Renals

AN INVESTIGATION OF USING HYBRID MODELING UNITS FOR IMPROVING END-TO-END SPEECH RECOGNITION SYSTEM ................................................................. 6743
Shunfei Chen, Xinhui Hu, Sheng Li, Xinkang Xu

FEDERATED ACOUSTIC MODELING FOR AUTOMATIC SPEECH RECOGNITION ...................... 6748
Xiaodong Cui, Songtao Lu, Brian Kingsbury

EAT: ENHANCED ASR-TTS FOR SELF-SUPERVISED SPEECH RECOGNITION ....................... 6753
Murali Karthick Baskar, Lukáš Burget, Shinji Watanabe, Ramon Fernandez Astudillo, Jan Honza Černocký

NEURAL ARCHITECTURE SEARCH FOR LF-MM TRAINED TIME DELAY NEURAL NETWORKS .................................................................................................................. 6758
Shoukang Hu, Xurong Xie, Shansong Liu, Mingyu Cui, Mengzhe Geng, Xunying Liu, Helen Meng

HYPOTHESIS STITCHER FOR END-TO-END SPEAKER-ATTRIBUTED ASR ON LONG-FORM MULTI-TALKER RECORDINGS ................................................................. 6763
Xuankai Chang, Naoyuki Kanda, Yashesh Gaur, Xiaofei Wang, Zhong Meng, Takuya Yoshioka
SPE-40: SPEECH RECOGNITION 14: ACOUSTIC MODELING 2

ENSEMBLE COMBINATION BETWEEN DIFFERENT TIME SEGMENTATIONS .......................................................... 6768
Jeremy H. M. Wong, Dimitrios Dimitriadis, Kenichi Kumatani, Yashesh Gaur, George
Polovets, Partha Parthasarathy, Eric Sun, Jinyu Li, Yifan Gong

STREAMING END-TO-END SPEECH RECOGNITION WITH JOINTLY TRAINED NEURAL
FEATURE ENHANCEMENT .............................................................................................................................. 6773
Chanwoo Kim, Abhinav Garg, Dhananjaya Gowda, Seongkyu Mun, Changwoo Han

TRANSFORMER IN ACTION: A COMPARATIVE STUDY OF TRANSFORMER-BASED
ACOUSTIC MODELS FOR LARGE SCALE SPEECH RECOGNITION APPLICATIONS ............................. 6778
Yongqiang Wang, Yangyang Shi, Frank Zhang, Chunyang Wu, Julian Chan, Ching-Feng Yeh,
Alex Xiao

EMFORMER: EFFICIENT MEMORY TRANSFORMER BASED ACOUSTIC MODEL FOR
LOW LATENCY STREAMING SPEECH RECOGNITION ............................................................................... 6783
Yangyang Shi, Yongqiang Wang, Chunyang Wu, Ching-Feng Yeh, Julian Chan, Frank Zhang,
Duc Le, Mike Seltzer

LEARNED TRANSFERABLE ARCHITECTURES CAN SURPASS HAND-DESIGNED
ARCHITECTURES FOR LARGE SCALE SPEECH RECOGNITION .......................................................... 6788
Liqiang He, Dan Su, Dong Yu

MULTITASK LEARNING AND JOINT OPTIMIZATION FOR TRANSFORMER-RNN-
TRANSUDER SPEECH RECOGNITION ................................................................................................. 6793
Jae-Jin Jeon, Eesung Kim

SPE-41: VOICE ACTIVITY AND DISFLUENCY DETECTION

SEP-28K: A DATASET FOR STUTTERING EVENT DETECTION FROM PODCASTS WITH
PEOPLE WHO STUTTER ......................................................................................................................... 6798
Colin Lea, Vikramjit Mitra, Aparna Joshi, Sachin Kajarekar, Jeffrey P. Bigham

A HYBRID CNN-BILSTM VOICE ACTIVITY DETECTOR ........................................................................ 6803
Nicholas Wilkinson, Thomas Niesler

SELF-ATTENTIVE VAD: CONTEXT-AWARE DETECTION OF VOICE FROM NOISE .......................... 6808
Yong Rae Jo, Young Ki Moon, Won Ik Cho, Geun Sik Jo

PREVENTING EARLY ENDPOINTING FOR ONLINE AUTOMATIC SPEECH
RECOGNITION ............................................................................................................................................ 6813
Yingzhu Zhao, Chongjia Ni, Cheung-Chi Leung, Shafiq Joty, Eng Siong Chng, Bin Ma

MARBLENET: DEEP 1D TIME-CHANNEL SEPARABLE CONVOLUTIONAL NEURAL
NETWORK FOR VOICE ACTIVITY DETECTION .................................................................................. 6818
Fei Jia, Someshbra Majumdar, Boris Ginsburg

SPEECH ENHANCEMENT AIDED END-TO-END MULTI-TASK LEARNING FOR VOICE
ACTIVITY DETECTION .......................................................................................................................... 6823
Xu Tan, Xiao-Lei Zhang

ROBUST VOICE ACTIVITY DETECTION USING A MASKED AUDITORY ENCODER
BASED CONVOLUTIONAL NEURAL NETWORK ................................................................................. 6828
Nan Li, Longbiao Wang, Masashi Unoki, Sheng Li, Rui Wang, Meng Ge, Jianwu Dang
SPE-42: KEYWORD SPOTTING

A STAGE MATCH FOR QUERY-BY-EXAMPLE SPOKEN TERM DETECTION BASED ON STRUCTURE INFORMATION OF QUERY ................................................................. 6833
Junyao Zhan, Qianhua He, Jianbin Su, Yaxiong Li

KNOWLEDGE TRANSFER FOR EFFICIENT ON-DEVICE FALSE TRIGGER MITIGATION .......... 6838
Dighe Dighe, Erik Marchi, Srikanth Vishnubhotla, Sachin Kajarekar, Devang Naik

PROGRESSIVE VOICE TRIGGER DETECTION: ACCURACY VS LATENCY .......................... 6843
Siddharth Sigtia, John Bridle, Hywel Richards, Pascal Clark, Erik Marchi, Vineet Garg

DYNAMIC CURRICULUM LEARNING VIA DATA PARAMETERS FOR NOISE ROBUST
KEYWORD SPOTTING ............................................................................................................................... 6848
Takuya Higuchi, Shreyas Saxena, Mehrez Souden, Tien Dung Tran, Masood Delfarah,
Chandra Dhir

CNN-BASED SPOKEN TERM DETECTION AND LOCALIZATION WITHOUT DYNAMIC
PROGRAMMING .......................................................................................................................................... 6853
Tzeviya Sylvia Fuchs, Yael Segal, Joseph Keshet

QUERY-BY-EXAMPLE KEYWORD SPOTTING SYSTEM USING MULTI-HEAD
ATTENTION AND SOFT-TRIPLE LOSS .................................................................................................... 6858
Jinmiao Huang, Waseem Gharbieh, Han Suk Shim, Eugene Kim

SPE-43: SPEECH RECOGNITION 15: ROBUST SPEECH RECOGNITION 1

A CLOSER LOOK AT AUDIO-VISUAL MULTI-PERSON SPEECH RECOGNITION AND
ACTIVE SPEAKER SELECTION ................................................................................................................ 6863
Otavio Braga, Olivier Siohan

GENERALIZED KNOWLEDGE DISTILLATION FROM AN ENSEMBLE OF SPECIALIZED
TEACHERS LEVERAGING UNSUPERVISED NEURAL CLUSTERING ............................................... 6868
Takashi Fukuda, Gakuto Kurata

MULTISTREAM CNN FOR ROBUST ACOUSTIC MODELING ......................................................... 6873
Kyu J. Han, Jing Pan, Venkata Krishna Naveen Tadala, Tao Ma, Dan Povey

IMPROVED ROBUSTNESS TO DISFLUENCIES IN RNN-TRANSDUCER BASED SPEECH
RECOGNITION ............................................................................................................................................. 6878
Valentin Mendelev, Tina Raissi, Guglielmo Camporese, Manuel Giollo

REPRESENTATION LEARNING FOR SPEECH RECOGNITION USING FEEDBACK BASED
RELEVANCE WEIGHTING ............................................................................................................................ 6883
Purvi Agrawal, Sriram Ganapathy

TOWARDS DATA SELECTION ON TTS DATA FOR CHILDREN’S SPEECH RECOGNITION .......... 6888
Wei Wang, Zhikai Zhou, Yizhou Lu, Hongji Wang, Chenpeng Du, Yanmin Qian

SPE-44: SPEECH RECOGNITION 16: ROBUST SPEECH RECOGNITION 2

AN INVESTIGATION OF END-TO-END MODELS FOR ROBUST SPEECH RECOGNITION .......... 6893
Archiki Prasad, Preethi Jyothi, Rajbabu Velmurugan
END-TO-END DEREVERBERATION, BEAMFORMING, AND SPEECH RECOGNITION WITH IMPROVED NUMERICAL STABILITY AND ADVANCED FRONTEND .......................................................... 6898
Wangyou Zhang, Christoph Boeddeker, Shinji Watanabe, Tomohiro Nakatani, Marc Delcroix, Keisuke Kinoshita, Tsubasa Ochiai, Naoyuki Kamo, Reinhold Haeb-Umbach, Yanmin Qian

STREAMING MULTI-SPEAKER ASR WITH RNN-T ................................................................. 6903
Ilya Sklyar, Anna Piunova, Yulan Liu

IMPROVING RNN TRANSUDER WITH TARGET SPEAKER EXTRACTION AND NEURAL UNCERTAINTY ESTIMATION ................................................................. 6908
Jiatong Shi, Chunlei Zhang, Chao Weng, Shinji Watanabe, Meng Yu, Dong Yu

A PROGRESSIVE LEARNING APPROACH TO ADAPTIVE NOISE AND SPEECH ESTIMATION FOR SPEECH ENHANCEMENT AND NOISY SPEECH RECOGNITION .................. 6913
Zhaoxu Nian, Yan-Hui Tu, Jun Du, Chin-Hui Lee

THE ACCENTED ENGLISH SPEECH RECOGNITION CHALLENGE 2020: OPEN DATASETS, TRACKS, BASELINES, RESULTS AND METHODS .......................................................... 6918
Xian Shi, Fan Yu, Yizhou Lu, Yuhaol Liang, Qiangze Feng, Dalin Wang, Yanmin Qian, Lei Xie

SPE-45: SPEECH ANALYSIS

COMPARATIVE STUDY OF DIFFERENT EPOCH EXTRACTION METHODS FOR SPEECH ASSOCIATED WITH VOICE DISORDERS ................................................................. 6923
Purva Barche, Krishna Gurugubelli, Anil Kumar Vuppala

HAVE YOU MADE A DECISION? WHERE? A PILOT STUDY ON INTERPRETABLE POLARITY ANALYSIS BASED ON ADVISING PROBLEM .......................................................... 6928
Tian Li, Jia-Chen Gu, Hui Liu, Quan Liu, Zhen-Hua Ling, Zhiming Su, Xiaodan Zhu

TRANSFORMER BASED UNSUPERVISED PRE-TRAINING FOR ACOUSTIC REPRESENTATION LEARNING ................................................................. 6933
Ruixiong Zhang, Haiwei Wu, Wubo Li, Dongwei Jiang, Wei Zou, Xiangang Li

A COMPARISON OF CONVOLUTIONAL NEURAL NETWORKS FOR GLOTTAL CLOSURE INSTANT DETECTION FROM RAW SPEECH ............................................................. 6938
Jindrich Matoušek, Daniel Tihelka

ENCODER-DECODER BASED PITCH TRACKING AND JOINT MODEL TRAINING FOR MANDARIN TONE CLASSIFICATION ................................................................. 6943
Hao Huang, Kai Wang, Ying Hu, Sheng Li

SPE-46: CORPORA AND OTHER RESOURCES

CONSTRUCTION OF A LARGE-SCALE JAPANESE ASR CORPUS ON TV RECORDINGS .......... 6948
Shintaro Ando, Hiromasa Fujihara

NISP: A MULTI-LINGUAL MULTI-ACCENT DATASET FOR SPEAKER PROFILING ............... 6953
Shareef Babu Kalluri, Deepu Vijayasenan, Sriram Ganapathy, Ragesh Rajan M, Prashant Krishnan

MULTILINGUAL PHONETIC DATASET FOR LOW RESOURCE SPEECH RECOGNITION ........... 6958
Xinjian Li, David R. Mortensen, Florian Metze, Alan W Black
AGE-VOX-Celeb: Multi-Modal Corpus for Facial and Speech Estimation
Naohiro Tawara, Atsunori Ogawa, Yuki Kitagishi, Hosana Kamiyama

DIDISPEECH: A Large Scale Mandarin Speech Corpus
Tingwei Guo, Cheng Wen, Dongwei Jiang, Ne Luo, Ruixiong Zhang, Shuaijiang Zhao, Wubu Li, Cheng Gong, Wei Zou, Kun Han, Xiangang Li

The In-The-Wild Speech Medical Corpus
Joana Correia, Francisco Teixeira, Catarina Botelho, Isabel Trancoso, Bhiksha Raj

SPE-47: Speech Recognition 17: Speech Adaptation and Normalization

Multiple-Hypothesis CTC-Based Semi-Supervised Adaptation of End-to-End Speech Recognition
Cong-Thanh Do, Rama Doddipatla, Thomas Hain

Vowel Non-Vowel Based Spectral Warping and Time Scale Modification for Improvement in Children’s ASR
Hemant Kathania, Avinash Kumar, Mikko Kurimo

Extending Parrotron: An End-to-End, Speech Conversion and Speech Recognition Model for Atypical Speech
Rohan Doshi, Youzheng Chen, Liyang Jiang, Xia Zhang, Fadi Biadsy, Bhuvana Ramabhadran, Fang Chu, Andrew Rosenberg, Pedro J. Moreno

Fundamental Frequency Feature Normalization and Data Augmentation for Child Speech Recognition
Gary Yeung, Ruchao Fan, Abeer Alwan

Analysis of X-Vectors for Low-Resource Speech Recognition
Martin Karafiát, Karel Veselý, Jan Honza Cernocký, Jan Profant, Jiri Nytra, Miroslav Hlaváček, Tomáš Pavlícek

Refining Automatic Speech Recognition System for Older Adults
Liu Chen, Meysam Asgari

SPE-48: Speech Recognition 18: Low Resource ASR

MIXSPEECH: Data Augmentation for Low-Resource Automatic Speech Recognition
Linghui Meng, Jin Xu, Xu Tan, Jindong Wang, Tao Qin, Bo Xu

End-to-End Multilingual Automatic Speech Recognition for Less-Resourced Languages: The Case of Four Ethiopian Languages
Solomon Teferra Abate, Martha Yifira Tachbelie, Tanja Schultz

Improved Data Selection for Domain Adaptation in ASR
Shannon Wotherspoon, William Hartmann, Matthew Snover, Owen Kimball

BI-APC: Bidirectional Autoregressive Predictive Coding for Unsupervised Pre-Training and Its Application to Children’s ASR
Ruchao Fan, Amber Afshan, Abeer Alwan

Meta-Adapter: Efficient Cross-Lingual Adaptation with Meta-Learning
Wenxin Hou, Yidong Wang, Shengzhou Gao, Takahiro Shinozaki
SPE-49: SPEECH SYNTHESIS 7: GENERAL TOPICS

CONTEXT-AWARE PROSODY CORRECTION FOR TEXT-BASED SPEECH EDITING
Max Morrison, Lucas Rencker, Zeyu Jin, Nicholas J. Bryan, Juan-Pablo Caceres, Bryan Pardo

FAST DCTTS: EFFICIENT DEEP CONVOLUTIONAL TEXT-TO-SPEECH
Minsu Kang, JiHyun Lee, Simin Kim, Injung Kim

SPEECH PREDICTION IN SILENT VIDEOS USING VARIATIONAL AUTOENCODERS
Ravindra Yadav, Ashish Sardana, Vinay P Namboodiri, Rajesh M Hegde

LEARNING DISENTANGLLED PHONE AND SPEAKER REPRESENTATIONS IN A SEMI-SUPERVISED VQ-VAE PARADIGM
Jennifer Williams, Yi Zhao, Erica Cooper, Junichi Yamagishi

HIGH-INTELLIGIBILITY SPEECH SYNTHESIS FOR DYSARTHRIC SPEAKERS WITH LPCNET-BASED TTS AND CYCLEVAE-BASED VC
Keisuke Matsubara, Takuma Okamoto, Ryoichi Takashima, Tetsuya Takiguchi, Tomoki Toda, Yoshinori Shiga, Hisashi Kawai

DENOISPEECH: DENOISING TEXT TO SPEECH WITH FRAME-LEVEL NOISE MODELING
Chen Zhang, Yi Ren, Xu Tan, Jinglin Liu, Kejun Zhang, Tao Qin, Sheng Zhao, Tie-Yan Liu

SPE-50: VOICE CONVERSION & SPEECH SYNTHESIS: SINGING VOICE & OTHER TOPICS

NON-AUTOREGRESSIVE SEQUENCE-TO-SEQUENCE VOICE CONVERSION
Tomoki Hayashi, Wen-Chin Huang, Kazuhiro Kobayashi, Tomoki Toda

PPG-BASED SINGING VOICE CONVERSION WITH ADVERSARIAL REPRESENTATION LEARNING
Zhonghao Li, Benlai Tang, Xiang Yin, Yuan Wan, Ling Xu, Chen Shen, Zejun Ma

LITESING: TOWARDS FAST, LIGHTWEIGHT AND EXPRESSIVE SINGING VOICE SYNTHESIS
Xiaobin Zhuang, Tao Jiang, Szu-Yu Chou, Bin Wu, Peng Hu, Simon Lui

SEMI-SUPERVISED LEARNING FOR SINGING SYNTHESIS TIMBRE
Jordi Bonada, Merlijn Blaauw

RECURRENT PHASE RECONSTRUCTION USING ESTIMATED PHASE DERIVATIVES FROM DEEP NEURAL NETWORKS
Lars Thieling, Daniel Wilhelm, Peter Jax

STABLE CHECKPOINT SELECTION AND EVALUATION IN SEQUENCE TO SEQUENCE SPEECH SYNTHESIS
Slava Shechtman, David Haws, Raul Fernandez
SPE-51: SPEECH ENHANCEMENT 7: SINGLE-CHANNEL PROCESSING

TSTNN: TWO-STAGE TRANSFORMER BASED NEURAL NETWORK FOR SPEECH ENHANCEMENT IN THE TIME DOMAIN.................................................................................................................. 7098
  Kai Wang, Bengbeng He, Wei-Ping Zhu

SELF-ATTENTION GENERATIVE ADVERSARIAL NETWORK FOR SPEECH ENHANCEMENT.......................................................................................................................... 7103
  Huy Phan, Huy Le Nguyen, Oliver Y. Chén, Philipp Koch, Ngọc Q. K. Duong, Ian McLaughlin, Alfred Mertins

NEURAL KALMAN FILTERING FOR SPEECH ENHANCEMENT .......................................................................................................................... 7108
  Wei Xue, Gang Quan, Chao Zhang, Guohong Ding, Xiaodong He, Bowen Zhou

NEURAL NOISE EMBEDDING FOR END-TO-END SPEECH ENHANCEMENT WITH CONDITIONAL LAYER NORMALIZATION .................................................................................................................. 7113
  Zhihui Zhang, Xiaoli Li, Yaxing Li, Yuanjie Dong, Dan Wang, Shengwu Xiong

PERCEPTUAL LOSS BASED SPEECH DENOISING WITH AN ENSEMBLE OF AUDIO PATTERN RECOGNITION AND SELF-SUPERVISED MODELS .................................................................................................................. 7118
  Saurabh Kataria, Jesús Villalba, Najim Dehak

TOWARDS AN ASR APPROACH USING ACOUSTIC AND LANGUAGE MODELS FOR SPEECH ENHANCEMENT .......................................................................................................................... 7123
  Khandokar Md. Nayem, Donald S. Williamson

SPE-52: SPEECH ENHANCEMENT 8: ECHO CANCELLATION AND OTHER TASKS

A NEURAL ACOUSTIC ECHO CANCELLER OPTIMIZED USING AN AUTOMATIC SPEECH RECOGNIZER AND LARGE SCALE SYNTHETIC DATA .................................................................................................. 7128
  Nathan Howard, Alex Park, Turaj Zakizadeh Shabestary, Alexander Gruenstein, Rohit Prabhavalkar

LOW-COMPLEXITY, REAL-TIME JOINT NEURAL ECHO CONTROL AND SPEECH ENHANCEMENT BASED ON PERCEPNET .................................................................................................................. 7133
  Jean-Marc Valin, Srikanth Tenneti, Karim Helwani, Umut Isik, Arvindh Krishnaswamy

ACOUSTIC ECHO CANCELLATION WITH THE DUAL-SIGNAL TRANSFORMATION LSTM NETWORK .......................................................................................................................... 7138
  Nils L. Westhausen, Bernd T. Meyer

HIGH FIDELITY SPEECH REGENERATION WITH APPLICATION TO SPEECH ENHANCEMENT .......................................................................................................................... 7143
  Adam Polyak, Lior Wolf, Yossi Adi, Ori Kabeli, Yaniv Taigman

A TIME-DOMAIN CONVOLUTIONAL RECURRENT NETWORK FOR PACKET LOSS CONCEALMENT .......................................................................................................................... 7148
  Ju Lin, Yun Wang, Kaustubh Kalganwala, Gil Keren, Didi Zhang, Christian Fuegen

CASCADED TIME + TIME-FREQUENCY UNET FOR SPEECH ENHANCEMENT: JOINTLY ADDRESSING CLIPPING, CODEC DISTORTIONS, AND GAPS .......................................................................................................................... 7153
  Arun Asokan Nair, Kazuhiro Koishida
**SPE-53: SPEAKER DIARIZATION**

HIDDEN MARKOV MODEL DIARISATION WITH SPEAKER LOCATION INFORMATION .......... 7158
Jeremy H. M. Wong, Xiong Xiao, Yifan Gong

COMPOSITIONAL EMBEDDING MODELS FOR SPEAKER IDENTIFICATION AND DIARISATION WITH SIMULTANEOUS SPEECH FROM 2+ SPEAKERS ........................................ 7163
Zeqian Li, Jacob Whitehill

CONTENT-AWARE SPEAKER EMBEDDINGS FOR SPEAKER DIARISATION ......................... 7168
G. Sun, D. Liu, C. Zhang, P. C. Woodland

MULTI-SCALE SPEAKER DIARISATION WITH NEURAL AFFINITY SCORE FUSION ............. 7173
Tae Jin Park, Manoj Kumar, Shrikanth Narayanan

A COMPARISON STUDY ON INFANT-PARENT VOICE DIARIZATION .................................. 7178
Junzhe Zhu, Mark Hasegawa-Johnson, Nancy L. McElwain

**SPE-54: END-TO-END SPEAKER DIARIZATION AND RECOGNITION**

END-TO-END DIARIZATION FOR VARIABLE NUMBER OF SPEAKERS WITH LOCAL-GLOBAL NETWORKS AND DISCRIMINATIVE SPEAKER EMBEDDINGS .................................. 7183
Soumi Maiti, Hakan Erdogan, Kevin Wilson, Scott Wisdom, Shinji Watanabe, John R. Hershey

END-TO-END SPEAKER DIARIZATION AS POST-PROCESSING ............................................. 7188
Shota Horiguchi, Paola García, Yusuke Fujita, Shinji Watanabe, Kenji Nagamatsu

BW-EDA-END: STREAMING END-TO-END NEURAL SPEAKER DIARIZATION FOR A VARIABLE NUMBER OF SPEAKERS ................................................................. 7193
Eunjung Han, Chul Lee, Andreas Stolcke

INTEGRATING END-TO-END NEURAL AND CLUSTERING-BASED DIARIZATION:
GETTING THE BEST OF BOTH WORLDS ............................................................................. 7198
Keisuke Kinoshita, Marc Delcroix, Naohiro Tawara

SIAMESE CAPSULE NETWORK FOR END-TO-END SPEAKER RECOGNITION IN THE WILD .............................................................. 7203
Amirhossein Hajavi, Ali Etemad

A REAL-TIME SPEAKER DIARIZATION SYSTEM BASED ON SPATIAL SPECTRUM .......... 7208
Siqi Zheng, Weilong Huang, Xianliang Wang, Hongbin Suo, Jinwei Feng, Zhijie Yan

**SPE-55: LANGUAGE IDENTIFICATION AND LOW RESOURCE SPEECH RECOGNITION**

UNSUPERVISED NEURAL ADAPTATION MODEL BASED ON OPTIMAL TRANSPORT FOR SPOKEN LANGUAGE IDENTIFICATION ............................................................. 7213
Xugang Lu, Peng Shen, Yu Tsao, Hisashi Kawai

JOINT ASR AND LANGUAGE IDENTIFICATION USING RNN-T: AN EFFICIENT APPROACH TO DYNAMIC LANGUAGE SWITCHING ....................................................... 7218
Surabhi Punjabi, Harish Ariskere, Zeynab Raeisy, Chander Chandak, Nikhil Bhave, Ankish Bansal, Markus Müller, Sergio Murillo, Ariya Rastrow, Andreas Stolcke, Jasha Droppo, Sri Garimella, Roland Maas, Mat Hans, Athanasios Mouchtaris, Siegfried Kunzmann
SPOKEN LANGUAGE IDENTIFICATION IN UNSEEN TARGET DOMAIN USING WITHIN-SAMPLE SIMILARITY LOSS.............................................................. 7223
H Muralikrishna, Shantanu Kapoor, Dileep Aroor Dinesh, Padmanabhan Rajan

EXPLORING THE USE OF COMMON LABEL SET TO IMPROVE SPEECH RECOGNITION OF LOW RESOURCE INDIAN LANGUAGES............................... 7228
Vishwas M. Shetty, S. Umesh

PHONE DISTRIBUTION ESTIMATION FOR LOW RESOURCE LANGUAGES.............................................. 7233
Xinjian Li, Juncheng Li, Jiali Yao, Alan W Black, Florian Metze

HOW PHONOTACTICS AFFECT MULTILINGUAL AND ZERO-SHOT ASR PERFORMANCE ................................................................. 7238
Siyuan Feng, Piotr Zelasko, Laureano Moro-Velázquez, Ali Abavisani, Mark Hasegawa-Johnson, Odette Scharenborg, Najim Dehak

SPE-56: PARALINGUISTICS IN SPEECH

MODELLING PARALINGUISTIC PROPERTIES IN CONVERSATIONAL SPEECH TO DETECT BIPOLAR DISORDER AND BORDERLINE PERSONALITY DISORDER ....... 7243
Bo Wang, Yue Wu, Nemanja Vaci, Maria Liakata, Terry Lyons, Kate E A Saunders

AN ATTENTION MODEL FOR HYPERNASALITY PREDICTION IN CHILDREN WITH CLEFT PALATE .............................................................. 7248
Vikram C. Mathad, Nancy Scherer, Kathy Chapman, Julie Liss, Visar Berisha

AN END-TO-END SPEECH ACCENT RECOGNITION METHOD BASED ON HYBRID CTC/ATTENTION TRANSFORMER ASR ......................................... 7253
Qiang Gao, Haiwei Wu, Yangqing Sun, Yitao Duan

MULTI-TASK ESTIMATION OF AGE AND COGNITIVE DECLINE FROM SPEECH.......................... 7258
Yilin Pan, Venkata Srikanth Nallanthighal, Daniel Blackburn, Heidi Christensen, Aki Härmä

DEEPEMOCLUSTER: A SEMI-SUPERVISED FRAMEWORK FOR LATENT CLUSTER REPRESENTATION OF SPEECH EMOTIONS ........................................... 7263
Wei-Cheng Lin, Kusha Sridhar, Carlos Busso

THE ROLE OF TASK AND ACOUSTIC SIMILARITY IN AUDIO TRANSFER LEARNING: INSIGHTS FROM THE SPEECH EMOTION RECOGNITION CASE .................. 7268
Andreas Triantafyllopoulos, Björn W. Schuller

SPE-57: SPEECH, DEPRESSION AND SLEEPINESS

SPEECH-BASED DEPRESSION PREDICTION USING ENCODER-WEIGHT-ONLY TRANSFER LEARNING AND A LARGE CORPUS ........................................ 7273
Amir Harati, Elizabeth Shriberg, Tomasz Rutowski, Piotr Chlebek, Yang Lu, Ricardo Oliveira

ESTIMATING SEVERITY OF DEPRESSION FROM ACOUSTIC FEATURES AND EMBEDDINGS OF NATURAL SPEECH .................................................. 7278
Sri Harsha Dumpala, Sheri Rempel, Katerina Dikaios, Mehrri Sajjadiyan, Rudolf Uher, Sageev Oore

AUTOMATIC ELICITATION COMPLIANCE FOR SHORT-DURATION SPEECH BASED DEPRESSION DETECTION .................................................. 7283
Brian Stasak, Zhaocheng Huang, Dale Joachim, Julien Epps
DEEP NEURAL NETWORK EMBEDDINGS FOR THE ESTIMATION OF THE DEGREE OF SLEEPINESS .................................................... 7288
José Vicente Egas-López, Gábor Gosztolya

PAUSE-ENCODED LANGUAGE MODELS FOR RECOGNITION OF ALZHEIMER’S DISEASE AND EMOTION.................................................. 7293
Jiahong Yuan, Xingyu Cai, Kenneth Church

END-2-END MODELING OF SPEECH AND GAIT FROM PATIENTS WITH PARKINSON’S DISEASE: COMPARISON BETWEEN HIGH QUALITY VS. SMARTPHONE DATA ........................................... 7298

SPE-58: DYSARTHRIC SPEECH PROCESSING
A SEQUENTIAL CONTRASTIVE LEARNING FRAMEWORK FOR ROBUST DYSARTHRIC SPEECH RECOGNITION ................................................................. 7303
Lidan Wu, Daoming Zong, Shiliang Sun, Jing Zhao

AUTOMATIC AND PERCEPTUAL DISCRIMINATION BETWEEN DYSARTHRIA, APRAXIA OF SPEECH, AND NEUROTYPICAL SPEECH .................................................. 7308
Ina Kodrasi, Michaela Pernon, Marina Laganaro, Hervé Bourlard

EFFECT OF NOISE AND MODEL COMPLEXITY ON DETECTION OF AMYOTROPHIC LATERAL SCLEROSIS AND PARKINSON’S DISEASE USING PITCH AND MFCC ................................................... 7313
Tanuka Bhattacharjee, Jhansi Mallela, Yamini Belur, Nalini Atchayaram, Ravi Yadav, Pradeep Reddy, Dipanjan Gope, Prasanta Kumar Ghosh

MULTI-TASK TRANSFORMER WITH INPUT FEATURE RECONSTRUCTION FOR DYSARTHRIC SPEECH RECOGNITION .................................................. 7318
Chaoyue Ding, Shiliang Sun, Jing Zhao

DETECTING ALZHEIMER’S DISEASE FROM SPEECH USING NEURAL NETWORKS WITH BOTTLENECK FEATURES AND DATA AUGMENTATION .................................................. 7323
Zhaoci Liu, Zhiqiang Guo, Zhenhua Ling, Yunxia Li

AUTOMATIC DYSARTHRIC SPEECH DETECTION EXPLOITING PAIRWISE DISTANCE-BASED CONVOLUTIONAL NEURAL NETWORKS .................................................. 7328
Parvaneh Janbakhshi, Ina Kodrasi, Hervé Bourlard

HLT-1: LANGUAGE MODELING 1: FUSION AND TRAINING FOR END-TO-END ASR
IMPROVED NEURAL LANGUAGE MODEL FUSION FOR STREAMING RECURRENT NEURAL NETWORK TRANSDUCER ................................................................. 7333
Suyoun Kim, Yuan Shangguan, Jay Mahadeokar, Antoine Bruguier, Christian Fuegen, Michael L. Seltzer, Duc Le

INTERNAL LANGUAGE MODEL TRAINING FOR DOMAIN-ADAPTIVE END-TO-END SPEECH RECOGNITION ................................................................. 7338
Zhong Meng, Naoyuki Kanda, Yashesh Gaur, Sarangarajan Parthasarathy, Eric Sun, Liang Lu, Xie Chen, Jinyu Li, Yifan Gong

SPEECH RECOGNITION BY SIMPLY FINE-TUNING BERT .................................................. 7343
Wen-Chin Huang, Chia-Hua Wu, Shang-Bao Luo, Kuan-Yu Chen, Hsin-Min Wang, Tomoki Toda
PERSONALIZATION STRATEGIES FOR END-TO-END SPEECH RECOGNITION SYSTEMS........ 7348
Aditya Gourav, Linda Liu, Ankur Gandhe, Yile Gu, Guitang Lan, Xiangyang Huang,
Shashank Kalmane, Gautam Tiwari, Denis Filimonov, Ariya Rastrow, Andreas Stolcke, Ivan
Bulyko

IMPROVING ENTITY RECALL IN AUTOMATIC SPEECH RECOGNITION WITH NEURAL
EMBEDDINGS.................................................................................................................. 7353
Christopher Li, Pat Rondon, Diamantino Caseiro, Leonid Velikovich, Xavier Velez, Petar
Aleksic

ADAPTABLE MULTI-DOMAIN LANGUAGE MODEL FOR TRANSFORMER ASR........... 7358
Taewoo Lee, Min-Joong Lee, Tae Gyoong Kang, Seokyeoung Jung, Minseok Kwon, Yeona
Hong, Jungin Lee, Kyoung-Gu Woo, Ho-Gyeong Kim, Jiseung Jeong, Jihyun Lee, Hosik Lee,
Young Sang Choi

HLT-2: LANGUAGE MODELING 2: NEURAL LANGUAGE MODELS

TRANSFORMER LANGUAGE MODELS WITH LSTM-BASED CROSS-UTTERANCE
INFORMATION REPRESENTATION ......................................................................................... 7363
G. Sun, C. Zhang, P. C. Woodland

LARGE MARGIN TRAINING IMPROVES LANGUAGE MODELS FOR ASR .................. 7368
Jilin Wang, Jiaji Huang, Kenneth Ward Church

DOMAIN-AWARE NEURAL LANGUAGE MODELS FOR SPEECH RECOGNITION........ 7373
Linda Liu, Yile Gu, Aditya Gourav, Ankur Gandhe, Shashank Kalmane, Denis Filimonov,
Ariya Rastrow, Ivan Bulyko

BAYESIAN TRANSFORMER LANGUAGE MODELS FOR SPEECH RECOGNITION........ 7378
Boyang Xue, Jianwei Yu, Junhao Xu, Shansong Liu, Shoukang Hu, Zi Ye, Mengzhe Geng,
Xunying Liu, Helen Meng

MIXED PRECISION QUANTIZATION OF TRANSFORMER LANGUAGE MODELS FOR
SPEECH RECOGNITION ...................................................................................................... 7383
Junhao Xu, Shoukang Hu, Jianwei Yu, Xunying Liu, Helen Meng

FEDERATED MARGINAL PERSONALIZATION FOR ASR RESCORING ....................... 7388
Zhe Liu, Fuchun Peng

HLT-3: DIALOGUE SYSTEMS 1: GENERAL TOPICS

MULTI PATH TRAINING FRAMEWORK FOR DATA-DRIVEN OPEN-DOMAIN
CONVERSATION SYSTEM ...................................................................................................... 7393
Sixing Wu, Dawei Zhang, Ying Li, Zhonghai Wu

ACTION STATE UPDATE APPROACH TO DIALOGUE MANAGEMENT .................. 7398
Svetlana Stoyanchev, Simon Keizer, Rama Doddipatla

GENERATING EMPATHETIC RESPONSES BY INJECTING ANTICIPATED EMOTION ....... 7403
Yuhan Liu, Jiachen Du, Xiang Li, Ruifeng Xu

TOWARDS IMMEDIATE BACKCHANNEL GENERATION USING ATTENTION-BASED
EARLY PREDICTION MODEL ............................................................................................. 7408
Amalia Istiqlali Adiba, Takeshi Homma, Toshinori Miyoshi
ERROR-DRIVEN PRUNING OF LANGUAGE MODELS FOR VIRTUAL ASSISTANTS
Sashank Gondala, Lyan Verwimp, Ernest Pusateri, Manos Tsagkias, Christophe Van Gysel

HLT-4: DIALOGUE SYSTEMS 2: RESPONSE GENERATION

PARAGRAPH LEVEL MULTI-PERSPECTIVE CONTEXT MODELING FOR QUESTION GENERATION
Jun Bai, Wenge Rong, Feiyu Xia, Yannmeng Wang, Yuanxin Ouyang, Zhang Xiong

IMPROVING DIALOGUE RESPONSE GENERATION VIA KNOWLEDGE GRAPH FILTER
Yannmeng Wang, Ye Wang, Xingyu Lou, Wenge Rong, Zhenghong Hao, Shaojun Wang

TOPIC-AWARE DIALOGUE GENERATION WITH TWO-HOP BASED GRAPH ATTENTION
Shijie Zhou, Wenge Rong, Jianfei Zhang, Yannmeng Wang, Libin Shi, Zhang Xiong

HSAN: A HIERARCHICAL SELF-ATTENTION NETWORK FOR MULTI-TURN DIALOGUE GENERATION
Yawei Kong, Lu Zhang, Can Ma, Cong Cao

LEARNING TO SELECT CONTEXT IN A HIERARCHICAL AND GLOBAL PERSPECTIVE FOR OPEN-DOMAIN DIALOGUE GENERATION
Lei Shen, Haolan Zhan, Xin Shen, Yang Feng

TOWARDS EFFICIENTLY DIVERSIFYING DIALOGUE GENERATION VIA EMBEDDING AUGMENTATION
Yu Cao, Liang Ding, Zhiliang Tian, Meng Fang

HLT-5: LANGUAGE UNDERSTANDING 1: END-TO-END SPEECH UNDERSTANDING

END2END ACOUSTIC TO SEMANTIC TRANSDUCTION
Valentin Pelloin, Nathalie Camelin, Antoine Laurent, Renato De Mori, Antoine Caubrière, Yannick Estève, Sylvain Meignier

ACOUSTICS BASED INTENT RECOGNITION USING DISCOVERED PHONETIC UNITS FOR LOW RESOURCE LANGUAGES
Akshat Gupta, Xinjian Li, Sai Krishna Rallabandi, Alan W Black

SPEECH-LANGUAGE PRE-TRAINING FOR END-TO-END SPOKEN LANGUAGE UNDERSTANDING
Yao Qian, Ximo Bian, Yu Shi, Naoyuki Kanda, Leo Shen, Zhen Xiao, Michael Zeng

TWO-STAGE TEXTUAL KNOWLEDGE DISTILLATION FOR END-TO-END SPOKEN LANGUAGE UNDERSTANDING
Seongbin Kim, Gyuwan Kim, Seongjin Shin, Sangmin Lee

SEMI-SUPERVISED SPOKEN LANGUAGE UNDERSTANDING VIA SELF-SUPERVISED SPEECH AND LANGUAGE MODEL PRETRAINING
Cheng-I Lai, Yung-Sung Chuang, Hung-Yi Lee, Shang-Wen Li, James Glass

DO AS I MEAN, NOT AS I SAY: SEQUENCE LOSS TRAINING FOR SPOKEN LANGUAGE UNDERSTANDING
Milind Rao, Pranav Dheram, Gautam Tiwari, Anirudh Raju, Jasha Droppo, Ariya Rastrow, Andreas Stolcke
HLT-6: LANGUAGE UNDERSTANDING 2: END-TO-END SPEECH UNDERSTANDING 2

ST-BERT: CROSS-MODAL LANGUAGE MODEL PRE-TRAINING FOR END-TO-END SPOKEN LANGUAGE UNDERSTANDING ................................................................. 7478
   Minjeong Kim, Gyuwan Kim, Sang-Woo Lee, Jung-Woo Ha

END-TO-END SPOKEN LANGUAGE UNDERSTANDING USING TRANSFORMER NETWORKS AND SELF-SUPERVISED PRE-TRAINED FEATURES ............................................. 7483
   Edmilson Morais, Hong-Kwang J. Kuo, Samuel Thomas, Zoltán Tüske, Brian Kingsbury

SENTIMENT INJECTED ITERATIVELY CO-INTERACTIVE NETWORK FOR SPOKEN LANGUAGE UNDERSTANDING ........................................................................ 7488
   Zhiqi Huang, Fenglin Liu, Peilin Zhou, Yuexian Zou

RNN TRANSDUCER MODELS FOR SPOKEN LANGUAGE UNDERSTANDING ................................................................. 7493
   Samuel Thomas, Hong-Kwang J. Kuo, George Saon, Zoltán Tüske, Brian Kingsbury, Gakuto Kurata, Zvi Kons, Ron Hoory

LEVERAGING ACOUSTIC AND LINGUISTIC EMBEDDINGS FROM PRETRAINED SPEECH AND LANGUAGE MODELS FOR INTENT CLASSIFICATION ........................................ 7498
   Bidisha Sharma, Maulik Madhavi, Haizhou Li

HLT-7: SPEECH TRANSLATION 1: MODELS

ORTHROS: NON-AUTOREGRESSIVE END-TO-END SPEECH TRANSLATION WITH DUAL-DECODER .................................................................................. 7503
   Hirofumi Inaguma, Yosuke Higuchi, Kevin Duh, Tatsuya Kawahara, Shinji Watanabe

CASCADED MODELS WITH CYCLIC FEEDBACK FOR DIRECT SPEECH TRANSLATION ........................................ 7508
   Tsz Kin Lam, Shigehiko Schamoni, Stefan Riezler

JOINTLY TRAINED TRANSFORMERS MODELS FOR SPOKEN LANGUAGE TRANSLATION ...................................................................................... 7513
   Hari Krishna Vydana, Martin Karafiát, Katerina Zmolikova, Lukáš Burget, Honza Cernocký

EFFICIENT USE OF END-TO-END DATA IN SPOKEN LANGUAGE PROCESSING ........................................ 7518
   Yiting Lu, Yu Wang, Mark J. F. Gales

STREAMING SIMULTANEOUS SPEECH TRANSLATION WITH AUGMENTED MEMORY TRANSFORMER ........................................................................... 7523
   Xutai Ma, Yongqiang Wang, Mohammad Javad Dousti, Philipp Koehn, Juan Pino

AN EMPIRICAL STUDY OF END-TO-END SIMULTANEOUS SPEECH TRANSLATION DECODING STRATEGIES ..................................................................... 7528
   Ha Nguyen, Yannick Estève, Laurent Besacier

HLT-8: SPEECH TRANSLATION 2: ASPECTS

MODELING HOMOPHONE NOISE FOR ROBUST NEURAL MACHINE TRANSLATION ........................................ 7533
   Wenjie Qin, Xiang Li, Yuhai Sun, Deyi Xiong, Jianwei Cui, Bin Wang

MACHINE TRANSLATION VERBOSITY CONTROL FOR AUTOMATIC DUBBING ...................................................................................... 7538
   Surafel M. Lakew, Marcello Federico, Yue Wang, Cuong Hoang, Yogesh Virkar, Roberto Barra-Chicote, Robert Enyedi
IMPROVEMENTS TO PROSODIC ALIGNMENT FOR AUTOMATIC DUBBING ................................. 7543
Yogesh Virkar, Marcello Federico, Robert Enyedi, Roberto Barra-Chicote

IMAGE-ASSISTED TRANSFORMER IN ZERO-RESOURCE MULTI-MODAL
TRANSLATION .............................................................................................................. 7548
Ping Huang, Shiliang Sun, Hao Yang

SENTENCE BOUNDARY AUGMENTATION FOR NEURAL MACHINE TRANSLATION
ROBUSTNESS ............................................................................................................... 7553
Daniel Li, Te I, Naveen Arivazhagan, Colin Cherry, Dirk Padfield

AN EMPIRICAL STUDY ON TASK-ORIENTED DIALOGUE TRANSLATION ............................ 7558
Siyou Liu

HLT-9: STYLE AND TEXT NORMALIZATION

MAPGN: MASKED POINTER-GENERATOR NETWORK FOR SEQUENCE-TO-SEQUENCE
PRE-TRAINING ............................................................................................................. 7563
Mana Ihori, Naoki Makishima, Tomohiro Tanaka, Akihiko Takashima, Shota Orihashi, Ryo
Masumura

ALIGNING THE TRAINING AND EVALUATION OF UNSUPERVISED TEXT STYLE
TRANSFER ...................................................................................................................... 7568
Wanhui Qian, Fuqing Zhu, Jinzhu Yang, Jizhong Han, Songlin Hu

NEURAL INVERSE TEXT NORMALIZATION ........................................................................ 7573
Monica Sunkara, Chaitanya Shivade, Sravan Bodapati, Katrin Kirchhoff

GENERATING HUMAN READABLE TRANSCRIPT FOR AUTOMATIC SPEECH
RECOGNITION WITH PRE-TRAINED LANGUAGE MODEL .............................................. 7578
Junwei Liao, Yu Shi, Ming Gong, Linjun Shou, Sefik Eskimez, Liyang Lu, Hong Qu, Michael
Zeng

IMPROVING NEURAL TEXT NORMALIZATION WITH PARTIAL PARAMETER
GENERATOR AND POINTER-GENERATOR NETWORK .................................................... 7583
Weiwei Jiang, Junjie Li, Minchuan Chen, Jun Ma, Shaojun Wang, Jing Xiao

HLT-10: MULTI-MODALITY IN LANGUAGE

INCORPORATING SYNTACTIC AND PHONETIC INFORMATION INTO MULTIMODAL
WORD EMBEDDINGS USING GRAPH CONVOLUTIONAL NETWORKS ................................. 7588
Wenhao Zhu, Shuang Liu, Chaoming Liu

LIFI: TOWARDS LINGUISTICALLY INFORMED FRAME INTERPOLATION .......................... 7593
Aradhya Neeraj Mathur, Devansh Batra, Yaman Kumar Singla, Rajiv Rain Shah, Changyou
Chen, Roger Zimmermann

TRIPLE SEQUENCE GENERATIVE ADVERSARIAL NETS FOR UNSUPERVISED IMAGE
CAPTIONING .................................................................................................................. 7598
Yucheng Zhou, Wei Tao, Wenchang Zhang

ALIGN OR ATTEND? TOWARD MORE EFFICIENT AND ACCURATE SPOKEN WORD
DISCOVERY USING SPEECH-TO-IMAGE RETRIEVAL ................................................... 7603
Liming Wang, Xinsheng Wang, Mark Hasegawa-Johnson, Odette Scharenborg, Najim Dehak
TOWARDS PRACTICAL LIPREADING WITH DISTILLED AND EFFICIENT MODELS
Pingchuan Ma, Brais Martinez, Stavros Petridis, Maja Pantic

END-TO-END AUDIO-VISUAL SPEECH RECOGNITION WITH CONFORMERS
Pingchuan Ma, Stavros Petridis, Maja Pantic

HLT-11: LANGUAGE UNDERSTANDING 3: SPEECH UNDERSTANDING - GENERAL TOPICS

ASR N-BEST FUSION NETS
Xinyue Liu, Mingda Li, Luoxin Chen, Prashan Wanigasekara, Weitong Ruan, Haidar Khan, Wael Hamza, Chengwei Su

BOOSTING LOW-RESOURCE INTENT DETECTION WITH IN-SCOPE PROTOTYPICAL NETWORKS
Hongzhan Lin, Yuanmeng Yan, Guang Chen

CONVERSATIONAL QUERY REWRITING WITH SELF-SUPERVISED LEARNING
Hang Liu, Meng Chen, Youzheng Wu, Xiaodong He, Bowen Zhou

HANDLING CLASS IMBALANCE IN LOW-RESOURCE DIALOGUE SYSTEMS BY COMBINING FEW-SHOT CLASSIFICATION AND INTERPOLATION
Vishal Sunder, Eric Foster-Lussier

IMPROVING CROSS-DOMAIN SLOT FILLING WITH COMMON SYNTACTIC STRUCTURE
Luchen Liu, Xixun Lin, Peng Zhang, Bin Wang

JOINT INTENT DETECTION AND SLOT FILLING BASED ON CONTINUAL LEARNING MODEL
Yanfei Hui, Jianzong Wang, Ning Cheng, Fengying Yu, Tianbo Wu, Jing Xiao

HLT-12: LANGUAGE UNDERSTANDING 4: SEMANTIC UNDERSTANDING

KNOWLEDGE-BASED CHAT DETECTION WITH FALSE MENTION DISCRIMINATION
Wei Liu, Peijie Huang, Dongzhu Liang, Zihao Zhou

REPLACING HUMAN AUDIO WITH SYNTHETIC AUDIO FOR ON-DEVICE UNSPOKEN PUNCTUATION PREDICTION
Daria Soboleva, Ondrej Skopek, Márkus Šajgalík, Victor Carbune, Felix Weissenberger, Julia Proskurnia, Bogdan Prisacari, Daniel Valcarce, Justin Lu, Rohit Prabhavalkar, Balint Miklos

ADVERSARIAL GENERATIVE DISTANCE-BASED CLASSIFIER FOR ROBUST OUT-OF-DOMAIN DETECTION
Zhiyuan Zeng, Hong Xu, Keqing He, Yuanmeng Yan, Sihong Liu, Zijun Liu, Weiran Xu

GAN-BASED OUT-OF-DOMAIN DETECTION USING BOTH IN-DOMAIN AND OUT-OF-DOMAIN SAMPLES
Chaojie Liang, Peijie Huang, Wenbin Lai, Ziheng Ruan

PROGRESSIVE DIALOGUE STATE TRACKING FOR MULTI-DOMAIN DIALOGUE SYSTEMS
Jiahao Wang, Mingqian Liu, Xiaojun Quan
MULTI-STEP SPOKEN LANGUAGE UNDERSTANDING SYSTEM BASED ON ADVERSARIAL LEARNING .............................................................. 7673
  Yu Wang, Yilin Shen, Hongxia Jin

HLT-13: INFORMATION EXTRACTION

MULTI-ENTITY COLLABORATIVE RELATION EXTRACTION .......................................................... 7678
  Haozhuang Liu, Ziran Li, Dongming Sheng, Hai-Tao Zheng, Ying Shen

MULTI-GRANULARITY HETEROGENEOUS GRAPH FOR DOCUMENT-LEVEL RELATION EXTRACTION .......................................................... 7683
  Hengzhu Tang, Yanan Cao, Zhenyu Zhang, Ruipeng Jia, Fang Fang, Shi Wang

IMPROVING EVENT DETECTION BY EXPLOITING LABEL HIERARCHY ........................................... 7688
  Xiangyu Xi, Wei Ye, Tong Zhang, Quanxiu Wang, Shikun Zhang, Huixing Jiang, Wei Wu

IMPROVING NER IN SOCIAL MEDIA VIA ENTITY TYPE-COMPATIBLE UNKNOWN WORD SUBSTITUTION .......................................................... 7693
  Jian Xie, Kai Zhang, Lin Sun, Yindu Su, Chenxiang Xu

MORE: A METRIC LEARNING BASED FRAMEWORK FOR OPEN-DOMAIN RELATION EXTRACTION .......................................................... 7698
  Yutong Wang, Renze Lou, Kai Zhang, Mao Yan Chen, Yujiu Yang

"YOU SHOULD PROBABLY READ THIS": HEDGE DETECTION IN TEXT ........................................... 7703
  Denys Katerenchuk, Rivka Levitan

HLT-14: LANGUAGE REPRESENTATIONS

ENHANCING MODEL ROBUSTNESS BY INCORPORATING ADVERSARIAL KNOWLEDGE INTO SEMANTIC REPRESENTATION .......................................................... 7708
  Jinfeng Li, Tianyu Du, Xiangyu Liu, Rong Zhang, Hui Xue, Shouling Ji

ELBERT: FAST ALBERT WITH CONFIDENCE-WINDOW BASED EARLY EXIT ........................................... 7713
  Keli Xie, Siyuan Lu, Meiqi Wang, Zhongfeng Wang

DUALFORMER: A UNIFIED BIDIRECTIONAL SEQUENCE-TO-SEQUENCE LEARNING .......................................................... 7718
  Jen-Tzung Chien, Wei-Hsiang Chang

TASK AWARE MULTI-TASK LEARNING FOR SPEECH TO TEXT TASKS ........................................... 7723
  Sathish Indurthi, Mohd Abbas Zaidi, Nikhil Kumar Lakumarapu, Beomseok Lee, Hyojung Han, Seokchan Ahn, Sangha Kim, Chanwoo Kim, Inchul Hwang

LABEL-AWARE TEXT REPRESENTATION FOR MULTI-LABEL TEXT CLASSIFICATION .......................................................... 7728
  Hao Guo, Xiangyang Li, Lei Zhang, Jia Liu, Wei Chen

MIXUP REGULARIZED ADVERSARIAL NETWORKS FOR MULTI-DOMAIN TEXT CLASSIFICATION .......................................................... 7733
  Yuan Wu, Diana Inkpen, Ahmed El-Roby
HLT-15: LANGUAGE ASSESSMENT

MISPRONUNCIATION DETECTION IN NON-NATIVE (L2) ENGLISH WITH UNCERTAINTY MODELING ......................................................... 7738
Daniel Korzekwa, Jaime Lorenzo-Trueba, Szymon Zaporowski, Shira Calamaro, Thomas Drugman, Bozena Kostek

ATTENTION-BASED MULTI-ENCODER AUTOMATIC PRONUNCIATION ASSESSMENT .......... 7743
Binghuai Lin, Liyuan Wang

IMPROVING PRONUNCIATION ASSESSMENT VIA ORDINAL REGRESSION WITH ANCHORED REFERENCE SAMPLES ......................................................... 7748
Bin Su, Shaoguang Mao, Frank Soong, Yan Xia, Jonathan Tien, Zhiyong Wu

ANALYSING BIAS IN SPOKEN LANGUAGE ASSESSMENT USING CONCEPT ACTIVATION VECTORS .................................................................................. 7753
Xizi Wei, Mark J. F. Gales, Kate M. Knill

SENONE-AWARE ADVERSARIAL MULTI-TASK TRAINING FOR UNSUPERVISED CHILD TO ADULT SPEECH ADAPTATION ......................................................... 7758
Richeng Duan, Nancy F. Chen

CLASSIFYING SPEECH INTELLIGIBILITY LEVELS OF CHILDREN IN TWO CONTINUOUS SPEECH STYLES .................................................................................. 7763
Yeh-Sheng Lin, Shu-Chuan Tseng

HLT-16: APPLICATIONS IN NATURAL LANGUAGE

RECENT ADVANCES IN ARABIC SYNTACTIC DIACRITICS RESTORATION ................................................................. 7768
Yasser Hifny

MAKING PUNCTUATION RESTORATION ROBUST AND FAST WITH MULTI-TASK LEARNING AND KNOWLEDGE DISTILLATION ......................................................... 7773
Michael Hentschel, Emiru Tsunoo, Takao Okuda

VARIATIONAL DIALOGUE GENERATION WITH NORMALIZING FLOWS .................................................................................. 7778
Tien-Ching Luo, Jen-Tzung Chien

NN-KOG2P: A NOVEL GRAPHEME-TO-PHONEME MODEL FOR KOREAN LANGUAGE ........ 7783
Hwa-Yeon Kim, Jong-Hwan Kim, Jae-Min Kim

JOINT ALIGNMENT LEARNING-ATTENTION BASED MODEL FOR GRAPHEME-TO-PHONEME CONVERSION ........................................................................ 7788
Yonghe Wang, Feilong Bao, Hui Zhang, Guanglai Gao

HLT-17: LANGUAGE UNDERSTANDING 5: QUESTION ANSWERING AND READING COMPREHENSION

KNOWLEDGE DISTILLATION FOR IMPROVED ACCURACY IN SPOKEN QUESTION ANSWERING .................................................................................. 7793
Chenyu You, Nuo Chen, Yuexian Zou
COARSE-TO-CAREFUL: SEEKING SEMANTIC-RELATED KNOWLEDGE FOR OPEN-DOMAIN COMMONSENSE QUESTION ANSWERING
Luxi Xing, Yue Hu, Jing Yu, Yuqiang Xie, Wei Peng

LANGUAGE MODEL IS ALL YOU NEED: NATURAL LANGUAGE UNDERSTANDING AS QUESTION ANSWERING
Mahdi Namazifar, Alexandros Papangelis, Gokhan Tur, Dilek Hakkani-Tür

INTEGRATING SUBGRAPH-AWARE RELATION AND DIRECTION REASONING FOR QUESTION ANSWERING
Xu Wang, Shuai Zhao, Bo Cheng, Jiale Han, Li Yingting, Hao Yang, Ivan Sekulic, Guoshun Nan

ROLE AWARE MULTI-PARTY DIALOGUE QUESTION ANSWERING
Jui-Heng Hsu, Po-Wei Shen, Hung-Ting Su, Chen-Hsi Chang, Jia-Fong Yeh, Winston H. Hsu

MCR-NET: A MULTI-STEP CO-INTERACTIVE RELATION NETWORK FOR UNANSWERABLE QUESTIONS ON MACHINE READING COMPREHENSION
Wei Peng, Yue Hu, Jing Yu, Luxi Xing, Yuqiang Xie, Zihao Zhu, Yajing Sun

HLT-18: LANGUAGE UNDERSTANDING 6: SUMMARIZATION AND COMPREHENSION

HIERARCHICAL SPEAKER-AWARE SEQUENCE-TO-SEQUENCE MODEL FOR DIALOGUE SUMMARIZATION
Yuejie Lei, Yuanmeng Yan, Zhiyuan Zeng, Keqing He, Ximing Zhang, Weiran Xu

A LARGE-SCALE CHINESE LONG-TEXT EXTRACTIVE SUMMARIZATION CORPUS
Kai Chen, Guanyu Fu, Qingcai Chen, Baotian Hu

ADAPTIVE BI-DIRECTIONAL ATTENTION: EXPLORING MULTI-GRANULARITY REPRESENTATIONS FOR MACHINE READING COMPREHENSION
Nuo Chen, Fenglin Liu, Chenyu You, Peilin Zhou, Yuexian Zou

GRAPH ATTENTION AND INTERACTION NETWORK WITH MULTI-TASK LEARNING FOR FACT VERIFICATION
Rui Yang, Runze Wang, Zhen-Hua Ling

ENHANCING DEEP PARAPHRASE IDENTIFICATION VIA LEVERAGING WORD ALIGNMENT INFORMATION
Boxin Li, Tingwen Liu, Bin Wang, Lihong Wang

AN END-TO-END ACTOR-CRITIC-BASED NEURAL COREFERENCE RESOLUTION SYSTEM
Yu Wang, Yilin Shen, Hongxia Jin

ASPS-1: ARCHITECTURES

REDUCED-COMPLEXITY MODULAR POLYNOMIAL MULTIPLICATION FOR R-LWE CRYPTOSYSTEMS
Xinmiao Zhang, Keshab K. Parhi

SEIZURE DETECTION USING POWER SPECTRAL DENSITY VIA HYPERDIMENSIONAL COMPUTING
Lulu Ge, Keshab K. Parhi
FPGA HARDWARE DESIGN FOR PLENOPTIC 3D IMAGE PROCESSING ALGORITHM
TARGETING A MOBILE APPLICATION ................................................................. 7863
  Faraz Bhatti, Thomas Greiner

SLAP: A SPLIT LATENCY ADAPTIVE VLIW PIPELINE ARCHITECTURE WHICH
ENABLES ON-THE-FLY VARIABLE SIMD VECTOR-LENGTH ........................................ 7868
  Ashish Shrivastava, Alan Gatherer, Tong Sun, Sushma Wokhlu, Alex Chandra

UNSUPERVISED CLUSTERING OF TIME SERIES SIGNALS USING NEUROMORPHIC
ENERGY-EFFICIENT TEMPORAL NEURAL NETWORKS .................................................... 7873
  Shreyas Chaudhari, Harideep Nair, José M. F. Moura, John Paul Shen

ANGLE–OF–ARRIVAL (AOA) FACTORIZATION IN MULTIPATH CHANNELS .................... 7878
  Yu-Lin Wei, Romit Roy Choudhury

ASPS-2: ALGORITHM/ARCHITECTURE CO-DESIGN

SCALED FAST NESTED KEY EQUATION SOLVER FOR GENERALIZED INTEGRATED
INTERLEAVED BCH DECODERS .................................................................................... 7883
  Zhenshan Xie, Xinmiao Zhang

JOINT OPTIMIZATION FOR FULL-DUPLEX CELLULAR COMMUNICATIONS VIA
INTELLIGENT REFLECTING SURFACE ........................................................................ 7888
  Zhangjie Peng, Cunhua Pan, Zhenkun Zhang, Xianzhe Chen, Li Li, A. Lee Swindlehurst

A COLOR DOPPLER PROCESSING ENGINE WITH AN ADAPTIVE CLUTTER FILTER FOR
PORTABLE ULTRASOUND IMAGING DEVICES ............................................................ 7893
  Yi-Lin Lo, Chia-Hsiang Yang

CONVOLUTIONAL NEURAL NETWORK-AIDED BIT-FLIPPING FOR BELIEF
PROPAGATION DECODING OF POLAR CODES ............................................................. 7898
  Chieh-Fang Teng, Andrew Kuan-Shiuan Ho, Chen-Hsi Derek Wu, Sin-Sheng Wong, An-Yeu
  Andy Wu

TAMING VOTING ALGORITHMS ON GPUS FOR AN EFFICIENT CONNECTED
COMPONENT ANALYSIS ALGORITHM .......................................................................... 7903
  Florian Lemaitre, Arthur Hennequin, Lionel Lacassagne

POSITNN: TRAINING DEEP NEURAL NETWORKS WITH MIXED LOW-PRECISION POSIT .... 7908
  Gonçalo Raposo, Pedro Tomás, Nuno Roma

ASPS-3: IOT

BLUETOOTH LOW ENERGY AND CNN-BASED ANGLE OF ARRIVAL LOCALIZATION
IN PRESENCE OF RAYLEIGH FADING ........................................................................... 7913
  Zohreh Hajjakhondi-Meybodi, Mohammad Salimibeni, Arash Mohammadi, Konstantinos N.
  Platanitis

ROBUST DEVICE-FREE PROXIMITY DETECTION USING WIFI ...................................... 7918
  Yuqian Hu, M. Zahid Ozturk, Feng Zhang, Beibei Wang, K. J. Ray Liu

ONLINE DYNAMIC WINDOW (ODW) ASSISTED 2-STAGE LSTM INDOOR
LOCALIZATION FOR SMART PHONES ......................................................................... 7923
  Mohammadamin Atashi, Arash Mohammadi
OPTIMAL TOA LOCALIZATION FOR MOVING SENSOR IN ASYMMETRIC NETWORK
Sihao Zhao, Xiao-Ping Zhang, Xiaowei Cui, Mingquan Lu

LOW COMPLEXITY SLM FOR OFDMA SYSTEM WITH IMPLICIT SIDE INFORMATION
Shicheng Hu, Miao Yang, Kai Kang, Hua Qian

REDUCED-COMPLEXITY CHANNEL ESTIMATION BY HIERARCHICAL INTERPOLATION EXPLOITING SPARSITY FOR MASSIVE MIMO SYSTEMS WITH UNIFORM RECTANGULAR ARRAY
Chi-Shiang Wang, Pei-Yun Tsai

ASPS-4: AUTONOMOUS SYSTEMS

TRAFFIC SPEED FORECASTING VIA SPATIO-TEMPORAL ATTENTIVE GRAPH ISOMORPHISM NETWORK
Qing Yang, Ting Zhong, Fan Zhou

INFERRING HIGH-RESOLUTIONAL URBAN FLOW WITH INTERNET OF MOBILE THINGS
Fan Zhou, Xin Jing, Liang Li, Ting Zhong

TRANSFER LEARNING FOR INPUT ESTIMATION OF VEHICLE SYSTEMS
Liam M. Cronin, Soheil Sadeghi Eshkevari, Debarshi Sen, Shamim N. Pakzad

IDENTIFICATION OF DEEP BREATH WHILE MOVING FORWARD BASED ON MULTIPLE BODY REGIONS AND GRAPH SIGNAL ANALYSIS
Yunlu Wang, Cheng Yang, Menghan Hu, Jian Zhang, Qingli Li, Guangtao Zhai, Xiao-Ping Zhang

MULTI-OBJECT TRACKING USING POISSON MULTI-BERNOULLI MIXTURE FILTERING FOR AUTONOMOUS VEHICLES
Su Pang, Hayder Radha

ADAPTIVE RF FINGERPRINT DECOMPOSITION IN MICRO UAV DETECTION BASED ON MACHINE LEARNING
Chengtao Xu, Fengyu He, Bowen Chen, Yushan Jiang, Houbing Song

ASPS-5: AUDIO & IMAGES

DEPRESSION DETECTION BY ANALYSING EYE MOVEMENTS ON EMOTIONAL IMAGES
Ruizhe Shen, Qi Zhan, Yu Wang, Huimin Ma

WEAKLY SUPERVISED PATCH LABEL INFERENCE NETWORK WITH IMAGE PYRAMID FOR PAVEMENT DISEASES RECOGNITION IN THE WILD
Guixin Huang, Sheng Huang, Luwen Huangfu, Dan Yang

A WIRELESS REFERENCE ACTIVE NOISE CONTROL HEADPHONE USING COHERENCE BASED SELECTION TECHNIQUE
Xiaoyi Shen, Dongyuan Shi, Woon-Seng Gan

HOW TO USE TIME INFORMATION EFFECTIVELY? COMBINING WITH TIME SHIFT MODULE FOR LIPREADING
Mingfeng Hao, Mutallip Mamut, Nurbiya Yadikar, Alimjan Aysa, Kurban Ubul
EXPLORING THE APPLICATION OF SYNTHETIC AUDIO IN TRAINING KEYWORD SPOTTERS
Andrew Werchniak, Roberto Barra Chicote, Yuriy Mishchenko, Jasha Droppo, Jeff Condal, Peng Liu, Anish Shah

GRAPH ENHANCED QUERY REWRITING FOR SPOKEN LANGUAGE UNDERSTANDING SYSTEM
Siyang Yuan, Saurabh Gupta, Xing Fan, Derek Liu, Yang Liu, Chenlei Guo

ASPS-6: SENSING & SENSOR PROCESSING

DEEP NEURAL NETWORK BASED COUGH DETECTION USING BED-MOUNTED ACCELEROMETER MEASUREMENTS
Madhurananda Pahar, Igor Miranda, Andreas Diacon, Thomas Niesler

RADIO FREQUENCY BASED HEART RATE VARIABILITY MONITORING
Fengyu Wang, Xiaolu Zeng, Chenshu Wu, Beibei Wang, K. J. Ray Liu

DISCRETE COSINE TRANSFORM BASED CAUSAL CONVOLUTIONAL NEURAL NETWORK FOR DRIFT COMPENSATION IN CHEMICAL SENSORS
Diaa Badawi, Agamyrat Agambayev, Sule Ozev, A. Enis Çetin

WIFI-BASED DEVICE-FREE GESTURE RECOGNITION THROUGH-THE-WALL
Sai Deepika Regani, Beibei Wang, K. J. Ray Liu

SOUND RECOVERY FROM RADIO SIGNALS
Muhammed Zahid Ozturk, Chenshu Wu, Beibei Wang, K. J. Ray Liu

FULLY-NEURAL APPROACH TO VEHICLE WEIGHING AND STRAIN PREDICTION ON BRIDGES USING WIRELESS ACCELEROMETERS
Takaya Kawakatsu, Kenro Aihara, Atsuhiro Takasu, Jun Adachi, Haoqi Wang, Tonomori Nagayama

ASPS-7: DATA SCIENCE & MACHINE LEARNING

END TO END LEARNING FOR CONVOLUTIVE MULTI-CHANNEL WIENER FILTERING
Masahito Togami

MAKF-SR: MULTI-AGENT ADAPTIVE KALMAN FILTERING-BASED SUCCESSOR REPRESENTATIONS
Mohammad Salimiben, Parvin Malekzadeh, Arash Mohammad, Petros Spachos, Konstantinos N. Plataniotis

VARIATION-STABLE FUSION FOR PPG-BASED BIOMETRIC SYSTEM
Dae Yon Hwang, Bilal Taha, Dimitrios Hatzinakos

IMPROVING STABILITY OF ADVERSARIAL LI-ION CELL USAGE DATA GENERATION USING GENERATIVE LATENT SPACE MODELLING
Subhankar Chattoraj, Savon Pratther, Souvik Pratther, Hubert Konik

SQWA: STOCHASTIC QUANTIZED WEIGHT AVERAGING FOR IMPROVING THE GENERALIZATION CAPABILITY OF LOW-PRECISION DEEP NEURAL NETWORKS
Sungho Shin, Yoonho Boo, Wonyong Sung
A QUANTITATIVE ANALYSIS OF THE ROBUSTNESS OF NEURAL NETWORKS FOR TABULAR DATA .......................................................... 8057
Kavya Gupta, Beatrice Pesquet-Popescu, Fateh Kaakai, Jean-Christophe Pesquet

SS-1: BEAMFORMING FOR INTELLIGENT SURFACES

SPATIAL EQUALIZATION BEFORE RECEPTION: RECONFIGURABLE INTELLIGENT SURFACES FOR MULTI-PATH MITIGATION .......................................................... 8062
Hongliang Zhang, Lingyang Song, Zhu Han, H. Vincent Poor

INTERFERENCE ANALYSIS IN RECONFIGURABLE INTELLIGENT SURFACE-ASSISTED MULTIPLE-INPUT MULTIPLE-OUTPUT SYSTEMS ........................................... 8067
Jiang Liu, Xuewen Qian, Marco Di Renzo

CODEBOOK DESIGN FOR DUAL-POLARIZED ULTRA-MASSIVE MIMO COMMUNICATIONS AT MILLIMETER WAVE AND TERAHERTZ BANDS ........................................ 8072
Shuai Nie, Ian F. Akyildiz

PERFORMANCE ANALYSIS OF SPATIAL AND FREQUENCY DOMAIN INDEX-MODULATED RECONFIGURABLE INTELLIGENT METASURFACES ............................................. 8077
John A. Hodge, Kumar Vijay Mishra, Brian M. Sadler, Amir I. Zaghloul

META-LEARNING FOR 6G COMMUNICATION NETWORKS WITH RECONFIGURABLE INTELLIGENT SURFACES ........................................................................ 8082
Minchae Jung, Walid Saad

SS-2: DEEP LEARNING METHODS FOR SOLVING LINEAR INVERSE PROBLEMS

MODEL-INSPIRED DEEP LEARNING FOR LIGHT-FIELD MICROSCOPY WITH APPLICATION TO NEURON LOCALIZATION .......................................................... 8087
Pingfan Song, Herman Verinaz Jadan, Carmel L. Howe, Peter Quicke, Amanda J. Foust, Pier Luigi Dragotti

TIME-VARYING GRAPH SIGNAL INPAINTING VIA UNROLLING NETWORKS ........................................... 8092
Siheng Chen, Yonina C. Eldar

DEEP LEARNING FOR LINEAR INVERSE PROBLEMS USING THE PLUG-AND-PLAY PRIORS FRAMEWORK .................................................................................. 8098
Wei Chen, David Wipf, Miguel Rodrigues

A PLUG-AND-PLAY DEEP IMAGE PRIOR ...................................................................................... 8103
Zhaodong Sun, Fabian Latorre, Thomas Sanchez, Volkan Cevher

MRI IMAGE RECOVERY USING DAMPED DENOISING VECTOR AMP ............................................. 8108
Subrata Sarkar, Rizwan Ahmad, Philip Schniter

OVERCOMING MEASUREMENT INCONSISTENCY IN DEEP LEARNING FOR LINEAR INVERSE PROBLEMS: APPLICATIONS IN MEDICAL IMAGING ...................................... 8113
Marija Vella, João F. C. Mota
SS-3: MACHINE LEARNING IN WIRELESS NETWORKS

SCALABLE REINFORCEMENT LEARNING FOR ROUTING IN AD-HOC NETWORKS BASED ON PHYSICAL-LAYER ATTRIBUTES ................................................................. 8118
  Wei Cui, Wei Yu

BLIND CARBON COPY ON DIRTY PAPER: SEAMLESS SPECTRUM UNDERLAY VIA CANONICAL CORRELATION ANALYSIS ................................................................. 8123
  Mohamed Salah Ibrahim, Nicholas D. Sidiropoulos

AN ACTOR-CRITIC REINFORCEMENT LEARNING APPROACH TO MINIMUM AGE OF INFORMATION SCHEDULING IN ENERGY HARVESTING NETWORKS ........................................ 8128
  Shiyang Leng, Aylin Yener

MOVING OBJECT CLASSIFICATION WITH A SUB-6 GHZ MASSIVE MIMO ARRAY USING REAL DATA ........................................................................................................ 8133
  B. R. Manoj, Guoda Tian, Sara Gunnarsson, Fredrik Tufvesson, Erik G. Larsson

OPTIMIZING COVERAGE AND CAPACITY IN CELLULAR NETWORKS USING MACHINE LEARNING ........................................................................................................ 8138
  Ryan M. Dreifuerst, Samuel Daulton, Yuchen Qian, Paul Varkey, Maximilian Balandat, Sanjay Kasturia, A

UNSUPERVISED LEARNING FOR ASYNCHRONOUS RESOURCE ALLOCATION IN AD-HOC WIRELESS NETWORKS ................................................................. 8143
  Zhiyang Wang, Mark Eisen, Alejandro Ribeiro

SS-4: DATA SCIENCE METHODS FOR COVID-19

TWO-STAGE ADAPTIVE POOLING WITH RT-QPCR FOR COVID-19 SCREENING ................................................................. 8148
  Anoosheh Heidarzadeh, Krishna Narayanan

POINT OF CARE IMAGE ANALYSIS FOR COVID-19 ........................................................................................................ 8153

AN IMPROVED DATA DRIVEN DYNAMIC SIRD MODEL FOR PREDICTIVE MONITORING OF COVID-19 .......................................................................................... 8158
  Pushpendra Singh, Amit Singhal, Binish Fatimah, Anubha Gupta

LEVERAGING A MULTIPLE-STRAIN MODEL WITH MUTATIONS IN ANALYZING THE SPREAD OF COVID-19 ....................................................................................... 8163
  Anirudh Sridhar, Osman Yagan, Rashad Eletreby, Simon A. Levin, Joshua B. Plotkin, H. Vincent Poor

CONTACT TRACING ENHANCES THE EFFICIENCY OF COVID-19 GROUP TESTING ................................................................................................................... 8168
  Ritesh Goenka, Shu-Jie Cao, Chau-Wai Wong, Ajit Rajwade, Dror Baron

OPTIMAL QUESTIONNAIRES FOR SCREENING OF STRATEGIC AGENTS ................................................................................ 8173
  Anuj S. Vora, Ankur A. Kulkarni
<table>
<thead>
<tr>
<th>SS-5: DOMAIN ADAPTATION FOR MULTIMEDIA SIGNAL PROCESSING</th>
</tr>
</thead>
</table>
| EXPLORING VISUAL-AUDIO COMPOSITION ALIGNMENT NETWORK FOR QUALITY FASHION RETRIEVAL IN VIDEO .................................................. 8178
  Yanhao Zhang, Jianmin Wu, Xiong Xiong, Dangwei Li, Chenwei Xie, Yun Zheng, Pan Pan, Yinghui Xu |
| A SECURE SEARCHABLE IMAGE RETRIEVAL SCHEME WITH CORRECT RETRIEVAL IDENTITY ................................................................. 8183
  Liejun Wang, Haitao Yu |
| INJECTING WORD INFORMATION WITH MULTI-LEVEL WORD ADAPTER FOR CHINESE SPOKEN LANGUAGE UNDERSTANDING .......................... 8188
  Dechuan Teng, Libo Qin, Wanxiang Che, Sendong Zhao, Ting Liu |
| A CO-INTERACTIVE TRANSFORMER FOR JOINT SLOT FILLING AND INTENT DETECTION ................................................................. 8193
  Libo Qin, Tailu Liu, Wanxiang Che, Bingbing Kang, Sendong Zhao, Ting Liu |
| DUAL METRIC DISCRIMINATOR FOR OPEN SET VIDEO DOMAIN ADAPTATION ............................................................... 8198
  Yatian Wang, Xiaolin Song, Yezhen Wang, Pengfei Xu, Runbo Hu, Hua Chai |
| CROSS-DOMAIN SENTIMENT CLASSIFICATION WITH CONTRASTIVE LEARNING AND MUTUAL INFORMATION MAXIMIZATION ..................... 8203
  Tian Li, Xiang Chen, Shanghang Zhang, Zhen Dong, Kurt Keutzer |

<table>
<thead>
<tr>
<th>SS-6: INTELLIGENT SENSING AND COMMUNICATIONS FOR EMERGING APPLICATIONS</th>
</tr>
</thead>
</table>
| LOW-COMPLEXITY PARAMETER LEARNING FOR OTFS MODULATION BASED AUTOMOTIVE RADAR ........................................................... 8208
  Chenwen Liu, Shengheng Liu, Zihuan Mao, Yongming Huang, Haiming Wang |
| FEDERATED DROPOUT LEARNING FOR HYBRID BEAMFORMING WITH SPATIAL PATH INDEX MODULATION IN MULTI-USER MMWAVE-MIMO SYSTEMS ................................................................. 8213
  Ahmet M. Elbir, Sinem Coleri, Kumar Vijay Mishra |
| INFORMATION DECODING AND SDR IMPLEMENTATION OF DFRC SYSTEMS WITHOUT TRAINING SIGNALS ................................................ 8218
  Daniel M. Wong, Batu K. Chalise, Justin Metcalf, Moeness Amin |
| A LOW-COMPLEXITY MIMO DUAL FUNCTION RADAR COMMUNICATION SYSTEM VIA ONE-BIT SAMPLING .............................................. 8223
  Siyu Zhu, Feng Xi, Shengyao Chen, Arye Nehorai |
| LEARNING TO SELECT FOR MIMO RADAR BASED ON HYBRID ANALOG-DIGITAL BEAMFORMING ...................................................... 8228
  Zhaoyi Xu, Fan Liu, Konstantinos Diamantaras, Christos Masouros, Athina Petropulu |
| WORD-LEVEL ASL RECOGNITION AND TRIGGER SIGN DETECTION WITH RF SENSORS ............................................................... 8233
  M. Mahbubur Rahman, Emre Kurtoglu, Robiulhossain Mdrafi, Ali C. Gurbuz, Evie Malaia, Chris Crawford, Darrin Griffin, Sevgi Z. Gurbuz |
SS-7: MULTI-FUNCTION RADIO FREQUENCY SYSTEM: RADAR, COMMUNICATION, POSITIONING AND BEYOND

HYBRID BEAMFORMING FOR WIDEBAND OFDM DUAL FUNCTION RADAR COMMUNICATIONS ........................................................................................................... 8238
Ziyang Cheng, Jinyang He, Shengnan Shi, Zishu He, Bin Liao

BIT CONSTRAINED COMMUNICATION RECEIVERS IN JOINT RADAR COMMUNICATIONS SYSTEMS ........................................................................................................... 8243
Dingyou Ma, Nir Shlezinger, Tianyao Huang, Yimin Liu, Yonina C. Eldar

ICI-AWARE PARAMETER ESTIMATION FOR MIMO-OFDM RADAR VIA APES SPATIAL FILTERING ........................................................................................................... 8248
Musa Furkan Keskin, Henk Wymeersch, Visa Koivunen

JOINT COMMUNICATIONS WITH FH-MIMO RADAR SYSTEMS: AN EXTENDED SIGNALING STRATEGY ........................................................................................................... 8253
Xiangrong Wang, Jing Xu, Aboulnasr Hassanien, Elias Aboutanios

FULL-DUPLEX MULTIFUNCTION TRANSCEIVER WITH JOINT CONSTANT ENVELOPE TRANSMISSION AND WIDEBAND RECEPTION ........................................................................................................... 8258
Jaakko Marin, Micael Bernhardt, Taneli Riihonen

WAVEFORM DESIGN FOR THE JOINT MIMO RADAR AND COMMUNICATIONS WITH LOW INTEGRATED SIDELOBE LEVELS AND ACCURATE INFORMATION EMBEDDING ........... 8263
Yongzhe Li, Xinyu Wu, Ran Tao

SS-8: NEAR-ML DECODING OF ERROR-CORRECTING CODES: ALGORITHMS AND IMPLEMENTATION

ORDERED RELIABILITY BITS GUESSING RANDOM ADDITIVE NOISE DECODING ............... 8268
Ken R. Duffy

LEARNED DECIMATION FOR NEURAL BELIEF PROPAGATION DECODERS ....................... 8273
Andreas Buchberger, Christian Häger, Henry D. Pfister, Laurent Schmalen, Alexandre Graell I Amat

ADMM-BASED ML DECODING: FROM THEORY TO PRACTICE ........................................... 8278
Kira Kraft, Norbert Wehn

TOWARDS PRACTICAL NEAR-MAXIMUM-LIKELIHOOD DECODING OF ERROR-CORRECTING CODES: AN OVERVIEW ................................................................. 8283
Thibaud Tonnellier, Marzieh Hashemipour, Nghia Doan, Warren J. Gross, Alexios Balatsoukas-Stimming

HIGH-THROUGHPUT VLSI ARCHITECTURE FOR SOFT-DECISION DECODING WITH ORBGRAND ........................................................................................................... 8288
Syed Mohsin Abbas, Thibaud Tonnellier, Furkan Ercan, Marwan Jalaleddine, Warren J. Gross

HARDWARE IMPLEMENTATION OF ITERATIVE PROJECTION-AGGREGATION DECODING OF REED-MULLER CODES ................................................................. 8293
Marzieh Hashemipour-Nazari, Kees Goossens, Alexios Balatsoukas-Stimming
SS-9: CONTACTLESS AND WIRELESS SENSING FOR SMART ENVIRONMENTS

M-ACTIVITY: ACCURATE AND REAL-TIME HUMAN ACTIVITY RECOGNITION VIA MILLIMETER WAVE RADAR ........................................................................................................................................  8298
Yuheng Wang, Haipeng Liu, Kening Cui, Anfu Zhou, Wensheng Li, Huadong Ma

PUSHING THE LIMIT OF PHASE OFFSET FOR CONTACTLESS SENSING USING COMMODITY WIFI .....................................................................................................................................  8303
Dongheng Zhang, Xiong Li, Yan Chen

NONCONTACT HEARTBEAT DETECTION BY VITERBI ALGORITHM WITH FUSION OF BEAT-BEAT INTERVAL AND DEEP LEARNING-DRIVEN BRANCH METRICS ....................................................................................................................  8308
Kohei Yamamoto, Tomoaki Ohtsuki

TYPINGWRISTBAND: A HUMAN SLIGHT MOTION SENSING SYSTEM BASED ON VIBRATION DETECTION ........................................................................................................................................  8313
Siyao Cheng, Jialiang Yan, Jianzhong Li, Jie Liu

MOVEMENT DETECTION USING A RECIPROCAL RECEIVED SIGNAL STRENGTH MODEL ..........................................................................................................................................................  8318
Ossi Kaltiokallio, Hüseyin Yigitler

DEEP CONVOLUTIONAL GAUSSIAN PROCESSES FOR MMWAVE OUTDOOR LOCALIZATION ..........................................................................................................................................  8323
Xuyu Wang, Mohini Patil, Chao Yang, Shiwen Mao, Palak Anilkumar Patel

SS-10: COMPUTER AUDITION FOR HEALTHCARE (CA4H)

EXPLORING AUTOMATIC COVID-19 DIAGNOSIS VIA VOICE AND SYMPTOMS FROM CROWDSOURCED DATA ........................................................................................................................................  8328
Jing Han, Chloë Brown, Jagmohan Chauhan, Andreas Grammenos, Apinan Hasthanasombat, Dimitris Spathis, Tong Xia, Pietro Cicuta, Cecilia Mascolo

COUGHWATCH: REAL-WORLD COUGH DETECTION USING SMARTWATCHES ........................................................................................................................................  8333
Daniyal Liaqat, Salaar Liaqat, Jun Lin Chen, Tina Sedaqhat, Moshe Gabel, Frank Rudzicz, Eyal De Lara

ACOUSTIC AND LINGUISTIC ANALYSES TO ASSESS EARLY-ONSET AND GENETIC ALZHEIMER’S DISEASE ........................................................................................................................................  8338

A NOISE-ROBUST SIGNAL PROCESSING STRATEGY FOR COCHLEAR IMPLANTS USING NEURAL NETWORKS ........................................................................................................................................  8343
Nengheng Zheng, Yupeng Shi, Yuyong Kang, Qinglin Meng

CONTEXT-AWARE SPEECH STRESS DETECTION IN HOSPITAL WORKERS USING BI-LSTM CLASSIFIERS ........................................................................................................................................  8348
Amr Gaballah, Abhishek Tiwari, Shrikanth Narayanan, Tiago H. Falk

UNSUPERVISED HEART ABNORMALITY DETECTION BASED ON PHONOCARDIOGRAM ANALYSIS WITH BETA VARIATIONAL AUTO-ENCODERS ........................................................................................................................................  8353
Shengchen Li, Ke Tian, Rui Wang
SS-11: ON-DEVICE AI FOR AUDIO AND SPEECH APPLICATIONS

COMPRESSING DEEP NEURAL NETWORKS FOR EFFICIENT SPEECH ENHANCEMENT .......... 8358
Ke Tan, Deliang Wang

IMPROVED MASK-CTC FOR NON-AUTOREGRESSIVE END-TO-END ASR ........................................ 8363
Yosuke Higuchi, Hirofumi Inaguma, Shinji Watanabe, Tetsuji Ogawa, Tetsunori Kobayashi

MEMORY-EFFICIENT SPEECH RECOGNITION ON SMART DEVICES .................................. 8368
Ganesh Venkatesh, Alagappan Valliappan, Jay Mahadeokar, Yuan Shangguan, Christian Fuegen, Michael L. Seltzer, Vikas Chandra

EXPEDITING DISCOVERY IN NEURAL ARCHITECTURE SEARCH BY COMBINING LEARNING WITH PLANNING ................................................................................................................... 8373
Farzaneh S. Fard, Vikrant Singh Tomar

SPECIALIZED EMBEDDING APPROXIMATION FOR EDGE INTELLIGENCE: A CASE STUDY IN URBAN SOUND CLASSIFICATION ................................................................. 8378
Sangeeta Srivastava, Dhrubojyoti Roy, Mark Cartwright, Juan P. Bello, Anish Arora

LIGHT-TTS: LIGHTWEIGHT MULTI-SPEAKER MULTI-LINGUAL TEXT-TO-SPEECH ................. 8383
Song Li, Beibei Ouyang, Lin Li, Qingyang Hong

SS-12: RECENT ADVANCES IN MMWAVE RADAR SENSING FOR AUTONOMOUS VEHICLES

EFFICIENT LONG PERIODIC BINARY SEQUENCE DESIGNS FOR AUTOMOTIVE RADAR ....... 8388
Yutao Chen, Ronghao Lin, Jian Li

JOINT LOCALIZATION AND PREDICTIVE BEAMFORMING IN VEHICULAR NETWORKS: POWER ALLOCATION BEYOND WATER-FILLING ...................................................... 8393
Fan Liu, Christos Masouros

A NEW AUTOMOTIVE RADAR 4D POINT CLOUDS DETECTOR BY USING DEEP LEARNING ........................................ 8398
Yuwei Cheng, Jingran Su, Hongyu Chen, Yimin Liu

ENHANCED AUTOMOTIVE TARGET DETECTION THROUGH RADAR AND COMMUNICATIONS SENSOR FUSION ............................................................................................................. 8403
Sayed Hossein Dokhanchi, Bhavani Shankar Mysore, Kumar Vijay Mishra, Björn Ottersten

EXTENDED OBJECT TRACKING WITH AUTOMOTIVE RADAR USING B-SPLINE CHAINED ELLIPSES MODEL ........................................................................................................ 8408
G. Yao, P. Wang, K. Berntorp, H. Mansour, P. Boufounos, P. V. Orlik

FOUR-DIMENSIONAL HIGH-RESOLUTION AUTOMOTIVE RADAR IMAGING EXPLOITING JOINT SPARSE-FREQUENCY AND SPARSE-ARRAY DESIGN ....................................... 8413
Shunqiao Sun, Yimn D. Zhang
SS-13: RECENT ADVANCES IN MULTICHANNEL AND MULTIMODAL MACHINE LEARNING FOR SPEECH APPLICATIONS

AN EMPIRICAL STUDY OF VISUAL FEATURES FOR DNN BASED AUDIO-VISUAL SPEECH ENHANCEMENT IN MULTI-TALKER ENVIRONMENTS......................................................... 8418
Shrishti Saha Shetu, Soumitro Chakrabarty, Emanuël A. P. Habets

ON THE ROLE OF VISUAL CUES IN AUDIOVISUAL SPEECH ENHANCEMENT................................. 8423
Zakaria Aldeneh, Anushree Prasanna Kumar, Barry-John Theobald, Erik Marchi, Sachin Kajarekar, Devang Naik, Ahmed Hussen Abdelaziz

CONVOLUTIVE TRANSFER FUNCTION INVARIANT SDR TRAINING CRITERIA FOR MULTI-CHANNEL REVERBERANT SPEECH SEPARATION.............................................................. 8428
Christoph Boeddeker, Wangyou Zhang, Tomohiro Nakatani, Keisuke Kinoshita, Tsubasa Ochiai, Marc Delcroix, Naoyuki Kamo, Yanmin Qian, Reinhold Haeb-Umbach

DIRECTIONAL ASR: A NEW PARADIGM FOR E2E MULTI-SPEAKER SPEECH RECOGNITION WITH SOURCE LOCALIZATION ......................................................................................... 8433
Aswin Shanmugam Subramanian, Chao Weng, Shinji Watanabe, Meng Yu, Yong Xu, Shixiong Zhang, Dong Yu

COMMUNICATION-COST AWARE MICROPHONE SELECTION FOR NEURAL SPEECH ENHANCEMENT WITH AD-HOC MICROPHONE ARRAYS......................................................... 8438
Jonah Casebeer, Jamshed Kaikaus, Paris Smaragdis

DEEP MULTI-FRAME MVDR FILTERING FOR SINGLE-MICROPHONE SPEECH ENHANCEMENT................................................................................................................................. 8443
Marvin Tammen, Simon Doclo

SS-14: ROBUST SENSING AND DETECTION IN CONGESTED SPECTRUM

COMPRESSIVE WIDEBAND SPECTRUM SENSING AND CARRIER FREQUENCY ESTIMATION WITH UNKNOWN MIMO CHANNELS................................................................. 8448
Hongwei Wang, Jilin Wang, Jun Fang, Hongbin Li

JOINT OPTIMIZATION OF SPECTRALLY CO-EXISTING MULTI-CARRIER RADAR AND COMMUNICATION SYSTEMS IN CLUTTERED ENVIRONMENTS.............................................. 8453
Fangzhou Wang, Hongbin Li, Braham Himed

TARGET DETECTION IN FREQUENCY HOPPING MIMO DUAL-FUNCTION RADAR-COMMUNICATION SYSTEMS ................................................................................................. 8458
Indu Priya Eedara, Moeness G. Amin, Giuseppe A. Fabrizio

ASYMPTOTIC DISTRIBUTION OF GENERALIZED LIKELIHOOD RATIO TEST UNDER MODEL MISSPECIFICATION WITH APPLICATION TO COOPERATIVE RADAR-COMMUNICATIONS......................................................... 8463
Akshay S. Bondre, Christ D. Richmond

ONLINE ANTENNA SELECTION FOR ENHANCED DOA ESTIMATION........................................ 8468
Elias Aboutanios, Hamed Nosrati, Xiangrong Wang

DESIGNING RANDOM FM RADAR WAVEFORMS WITH COMPACT SPECTRUM................................ 8473
Charles A. Mohr, Shannon D. Blunt
<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>SS-15: SIGNAL PROCESSING FOR COLLABORATIVE INTELLIGENCE</td>
<td></td>
</tr>
<tr>
<td>COLLABORATIVE INFERENCE VIA ENSEMBLES ON THE EDGE</td>
<td>8478</td>
</tr>
<tr>
<td>Nir Shlezinger, Erez Farhan, Hai Morgenstern, Yonina C. Eldar</td>
<td></td>
</tr>
<tr>
<td>ALLOCATING DNN LAYERS COMPUTATION BETWEEN FRONT-END DEVICES AND</td>
<td>8483</td>
</tr>
<tr>
<td>THE CLOUD SERVER FOR VIDEO BIG DATA PROCESSING</td>
<td></td>
</tr>
<tr>
<td>Peiyin Xing, Xiaofei Liu, Peixi Peng, Tiejun Huang, Yonghong Tian</td>
<td></td>
</tr>
<tr>
<td>BRANCHY-GNN: A DEVICE-EDGE CO-INFERENCE FRAMEWORK FOR EFFICIENT</td>
<td>8488</td>
</tr>
<tr>
<td>POINT CLOUD PROCESSING</td>
<td></td>
</tr>
<tr>
<td>Jiawei Shao, Haowei Zhang, Yuyi Mao, Jun Zhang</td>
<td></td>
</tr>
<tr>
<td>COLLABORATIVE INTELLIGENCE: CHALLENGES AND OPPORTUNITIES</td>
<td>8493</td>
</tr>
<tr>
<td>Ivan V. Bajic, Weisi Lin, Yonghong Tian</td>
<td></td>
</tr>
<tr>
<td>LATENT SPACE MOTION ANALYSIS FOR COLLABORATIVE INTELLIGENCE</td>
<td>8498</td>
</tr>
<tr>
<td>Mateen Ulhaq, Ivan V. Bajic</td>
<td></td>
</tr>
<tr>
<td>TEACHER- STUDENT LEARNING WITH MULTI-GRANULARITY CONSTRAINT</td>
<td>8503</td>
</tr>
<tr>
<td>TOWARDS COMPACT FACIAL FEATURE REPRESENTATION</td>
<td></td>
</tr>
<tr>
<td>Shurun Wang, Shiqi Wang, Wenhan Yang, Xinfeng Zhang, Shanshe Wang,</td>
<td></td>
</tr>
<tr>
<td>Siwei Ma</td>
<td></td>
</tr>
<tr>
<td>LEARNING THE RELEVANT SUBSTRUCTURES FOR TASKS ON GRAPH DATA</td>
<td>8528</td>
</tr>
<tr>
<td>Lei Chen, Zhengdiao Chen, Joan Bruna</td>
<td></td>
</tr>
<tr>
<td>A SHORT TUTORIAL ON THE WEISFEILER-LEHMAN TEST AND ITS VARIANTS</td>
<td>8533</td>
</tr>
<tr>
<td>Ningyuan Teresa Huang, Soledad Villar</td>
<td></td>
</tr>
<tr>
<td>CHLG-2: ZYELL - NCTUNETWORK ANOMALY DETECTION CHALLENGE</td>
<td></td>
</tr>
<tr>
<td>HYBRID MODEL FOR NETWORK ANOMALY DETECTION WITH GRADIENT BOOSTING</td>
<td>8538</td>
</tr>
<tr>
<td>DECISION TREES AND TABTRANSFORMER</td>
<td></td>
</tr>
<tr>
<td>Xinyue Xu, Xiaolu Zheng</td>
<td></td>
</tr>
<tr>
<td>VOTING-BASED ENSEMBLE MODEL FOR NETWORK ANOMALY DETECTION</td>
<td>8543</td>
</tr>
<tr>
<td>Tzu-Hsin Yang, Yu-Tai Lin, Chao-Lun Wu, Chih-Yu Wang</td>
<td></td>
</tr>
</tbody>
</table>
AN ACCURACY NETWORK ANOMALY DETECTION METHOD BASED ON ENSEMBLE MODEL .......................................................................................................................................................... 8548
    Fengrui Liu, Xuefei Li, Wei Xiong, Haiyang Jiang, Gaogang Xie

FDEN: MINING EFFECTIVE INFORMATION OF FEATURES IN DETECTING NETWORK ANOMALIES ................................................................................................................................................ 8553
    Bin Li, Yijie Wang, Mingyu Liu, Kele Xu, Zhongyang Wang, Li Cheng, Yizhou Li

CHLG-3: COVID-19 DIAGNOSIS

MULTI-SCALE RESIDUAL NETWORK FOR COVID-19 DIAGNOSIS USING CT-SCANS .............................................. 8558
    Pratyush Garg, Rishabh Ranjan, Kamini Upadhyay, Monika Agrawal, Desh Deepak

DIAGNOSING COVID-19 FROM CT IMAGES BASED ON AN ENSEMBLE LEARNING FRAMEWORK .............................................................................................................................................. 8563
    Bingyang Li, Qi Zhang, Yinan Song, Zhicheng Zhao, Zhu Meng, Fei Su

CNR-IEMN: A DEEP LEARNING BASED APPROACH TO RECOGNISE COVID-19 FROM CT-SCAN ....................................................................................................................................................... 8568
    Fares Bougourzi, Riccardo Contino, Cosimo Distante, Abdelmalik Taleb-Ahmed

COVID-19 DIAGNOSTIC USING 3D DEEP TRANSFER LEARNING FOR CLASSIFICATION OF VOLUMETRIC COMPUTERISED TOMOGRAPHY CHEST SCANS ........................................................................ 8573
    Shuohan Xue, Charith Abhayaratne

A MULTI-SCALE PROGRESSIVE LEARNING STRATEGY FOR COVID-19 DIAGNOSIS USING CHEST COMPUTED TOMOGRAPHY WITH IMBALANCED DATA ........................................................................ 8578
    Zaifeng Yang, Yubo Hou, Zhenghua Chen, Le Zhang, Jie Chen

DETECTING COVID-19 AND COMMUNITY ACQUIRED PNEUMONIA USING CHEST CT SCAN IMAGES WITH DEEP LEARNING ............................................................................................................... 8583
    Shubham Chaudhary, Sadbhawna Sadbhawna, Vinit Jakhetiya, Badri N Subudhi, Ujjwal Baid, Sharath Chandra Guntuku

CHLG-1: MULTI-SPEAKER MULTI-STYLE VOICE CLONING CHALLENGE (M2VOC)

INVESTIGATING ON INCORPORATING PRETRAINED AND LEARNABLE SPEAKER REPRESENTATIONS FOR MULTI-SPEAKER MULTI-STYLE TEXT-TO-SPEECH ........................................................................ 8588
    Chung-Ming Chien, Jheng-Hao Lin, Chien-Yu Huang, Po-Chun Hsu, Hung-Yi Lee

THE THINKIT SYSTEM FOR ICASSP2021 M2VOC CHALLENGE .......................................................................................................................... 8593
    Zengqiang Shang, Haozhe Zhang, Ziyi Chen, Bolin Zhou, Pengyuan Zhang

DIAN: DURATION INFORMED AUTO-REGRESSIVE NETWORK FOR VOICE CLONING .......................................................................................................................... 8598
    Wei Song, Xin Yuan, Zhengchen Zhang, Chao Zhang, Youzheng Wu, Xiaodong He, Bowen Zhou

PROSODY AND VOICE FACTORIZATION FOR FEW-SHOT SPEAKER ADAPTATION IN THE CHALLENGE M2VOC 2021 .......................................................................................................................................................... 8603
    Tao Wang, Ruibo Fu, Jiangyan Yi, Jianhua Tao, Zhengqi Wen, Chunyu Qiang, Shiming Wang

THE HUYA MULTI-SPEAKER AND MULTI-STYLE SPEECH SYNTHESIS SYSTEM FOR M2VOC CHALLENGE 2020 .......................................................................................................................................................... 8608
    Jie Wang, Yuren You, Feng Liu, Deyi Tuo, Shiying Kang, Zhiyong Wu, Helen Meng