2018 IEEE Spoken Language Technology Workshop (SLT 2018)

Athens, Greece 18-21 December 2018

Pages 1-541



IEEE Catalog Number: ISBN: CFP18SLT-POD 978-1-5386-4335-8

Copyright © 2018 by the Institute of Electrical and Electronics Engineers, Inc. All Rights Reserved

Copyright and Reprint Permissions: Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limit of U.S. copyright law for private use of patrons those articles in this volume that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923.

For other copying, reprint or republication permission, write to IEEE Copyrights Manager, IEEE Service Center, 445 Hoes Lane, Piscataway, NJ 08854. All rights reserved.

*** This is a print representation of what appears in the IEEE Digital Library. Some format issues inherent in the e-media version may also appear in this print version.

IEEE Catalog Number:	CFP18SLT-POD
ISBN (Print-On-Demand):	978-1-5386-4335-8
ISBN (Online):	978-1-5386-4334-1

Additional Copies of This Publication Are Available From:

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



TABLE OF CONTENTS

ASR I

#1: HIGH-DEGREE FEATURE FOR DEEP NEURAL NETWORK BASED ACOUSTIC
Hoon Chung, Sung Joo Lee, Jeon Gue Park, Electronics and Telecommunications Research Institute, Republic of Korea
#2: DENSENET BLSTM FOR ACOUSTIC MODELING IN ROBUST ASR
Maximilian Strake, Pascal Behr, Timo Lohrenz, Tim Fingscheidt, Technische Universität Braunschweig, Germany
#3: PHASE-BASED FEATURE REPRESENTATIONS FOR IMPROVING RECOGNITION13 OF DYSARTHRIC SPEECH
Siddharth Sehgal, Stuart Cunningham, Phil Green, University of Sheffield, United Kingdom
#4: EFFICIENT BUILDING STRATEGY WITH KNOWLEDGE DISTILLATION FOR
Yamaguchi, Manabu Okamoto, Yushi Aono, NTT Corporation, Japan
#5: ADVANCING MULTI-ACCENTED LSTM-CTC SPEECH RECOGNITION USING A29 DOMAIN SPECIFIC STUDENT-TEACHER LEARNING PARADIGM Shahram Ghorbani, Ahmet E. Bulut, John H.L. Hansen, University of Texas at Dallas, United States
#6: DYNAMIC EXTENSION OF ASR LEXICON USING WIKIPEDIA DATA
#7: IMPROVING LF-MMI USING UNCONSTRAINED SUPERVISIONS FOR ASR43 Hossein Hadian, Sharif University of Technology, Iran; Daniel Povey, Johns Hopkins University, United States; Hossein Sameti, Sharif University of Technology, Iran; Jan Trmal, Sanjeev Khudanpur, Johns Hopkins University, United States
#8: ON TRAINING RECURRENT NETWORKS WITH TRUNCATED
#9: LEARNING NOISE-INVARIANT REPRESENTATIONS FOR ROBUST SPEECH56 RECOGNITION Davis Liang, Zhiheng Huang, Amazon AI, United States; Zachary Lipton, Carnegie Mellon University, United States
#10: AN EXPLORATION OF DIRECTLY USING WORD AS ACOUSTIC MODELING64 UNIT FOR SPEECH RECOGNITION Chunlei Zhang, The University of Texas at Dallas, United States; Chengzhu Yu, Chao Weng, Jia Cui, Dong Yu, Tencent AI Lab, United States
#11: IMPROVED TRAINING OF NEURAL TRANS-DIMENSIONAL RANDOM FIELD

Bin Wang, Zhijian Ou, Tsinghua University, China

#12: IMPROVING VERY DEEP TIME-DELAY NEURAL NETWORK WITH77

VERTICAL-ATTENTION FOR EFFECTIVELY TRAINING CTC-BASED ASR SYSTEMS Sheng Li, Xugang Lu, Ryoichi Takashima, Peng Shen, National Institute of Information and Communications Technology, Japan; Tatsuya Kawahara, National Institute of Information and Communications Technology (NICT) / Kyoto University, Japan; Hisashi Kawai, National Institute of Information and Communications Technology, Japan

DETECTION, PARALINGUISTICS AND CODING

#1: EXPLORING END-TO-END ATTENTION-BASED NEURAL NETWORKS FOR84 NATIVE LANGUAGE IDENTIFICATION

Rutuja Ubale, Yao Qian, Keelan Evanini, Educational Testing Service Research, United States

Bhusan Chettri, Saumitra Mishra, Queen Mary University of London, United Kingdom; Bob L. Sturm, KTH Royal Institute of Engineering, Sweden; Emmanouil Benetos, Queen Mary University of London, United Kingdom

Xiaoxiao Miao, The University of Kent / Key Laboratory of Speech Acoustics and Content Understanding, Institute of Acoustics / University of Chinese Academy of Sciences, China; Ian McLoughlin, The University of Kent, United Kingdom; Shengyu Yao, Key Laboratory of Speech Acoustics and Content Understanding, Institute of Acoustics / University of Chinese Academy of Sciences, China; Yonghong Yan, Key Laboratory of Speech Acoustics and Content Understanding, Institute of Acoustics / University of Chinese Academy of Sciences / Xinjiang Key Laboratory of Minority Speech and Language Information Processing, Xinjiang Technical Institute of Physics and Chemistry, Chinese Academy of Sciences, China

#4: UNSUPERVISED REPRESENTATION LEARNING OF SPEECH FOR DIALECT105 IDENTIFICATION

Suwon Shon, Wei-Ning Hsu, James Glass, Massachusetts Institute of Technology, United States

#5: MULTIMODAL SPEECH EMOTION RECOGNITION USING AUDIO AND TEXT......112

Seunghyun Yoon, Seokhyun Byun, Kyomin Jung, Seoul National University, Republic of Korea

Gábor Gosztolya, MTA-SZTE Research Group on Artificial Intelligence, Hungary; Róbert Busa-Fekete, Yahoo Research Inc., United States

#7: CONTEXT-AWARE ATTENTION MECHANISM FOR SPEECH EMOTION126 RECOGNITION

Gaetan Ramet, Ecole Polytechnique Federale de Lausanne, Switzerland; Philip N. Garner, Idiap Research Institute, Switzerland; Michael Baeriswyl, Alexandros Lazaridis, Swisscom, Switzerland

#8: AN EXPERIMENTAL STUDY ON AUDIO REPLAY ATTACK DETECTION USING132 DEEP NEURAL NETWORKS

Bekir Bakar, Cemal Hanilci, Bursa Technical University, Turkey

#9: LSTM-BASED WHISPER DETECTION	9
Zeynab Raeesy, Kellen Gillespie, Chengyuan Ma, Thomas Drugman, Jiacheng Gu, Roland Maas, Ariya	
Rastrow, Björn Hoffmeister, Amazon, United States	
#10. A MEDICAN SICN I ANCHACE EINCEDSDELLING DECOGNITION IN THE WILD 14	5
#10: AMERICAN SIGN LANGUAGE FINGERSPELLING RECOGNITION IN THE WILD	3
Bowen Shi, Toyota Technological Institute at Chicago, United States; Aurora Martinez Del Rio, Jonathan	
Keane, University of Chicago, United States; Jonathan Michaux, Toyota Technological Institute at Chicago,	
United States; Diane Brentari, University of Chicago, United States; Greg Shakhnarovich, Karen Livescu,	
Toyota Technological Institute at Chicago, United States	
#11, WAVENET BASED ZEDO DEL AVI OSSI ESS SDEECH CODINC 15	2
#11: WAY ENE I-DASED LERO-DELAI LUSSLESS SPEECH CUDING	3

RECONSTRUCTION: FROM AUGMENTED ACOUSTIC INVERSION TO ARTICULATORY FEATURE RECONSTRUCTION WITHOUT ARTICULATORY DATA Rosanna Turrisi, Raffaele Tavarone, Leonardo Badino, Istituto Italiano di Tecnologia, Italy

#13: A DEEP LEARNING APPROACH FOR DATA DRIVEN VOCAL TRACT AREA167 FUNCTION ESTIMATION Sasan Asadiabadi, Engin Erzin, Koc university, Turkey

ASR II

#1: OCCAM'S ADAPTATION: A COMPARISON OF INTERPOLATION OF BASES174 ADAPTATION METHODS FOR MULTI-DIALECT ACOUSTIC MODELING WITH LSTMS *Mikaela Grace, Meysam Bastani, Eugene Weinstein, Google, United States*

#2: SPEECH CHAIN FOR SEMI-SUPERVISED LEARNING OF JAPANESE-ENGLISH182 CODE-SWITCHING ASR AND TTS Sahoko Nakayama Andros Tiandra Sakriani Sakti, Satoshi Nakamura, Nara Institute of Science and

Sahoko Nakayama, Andros Tjandra, Sakriani Sakti, Satoshi Nakamura, Nara Institute of Science and Technology, Japan

Minguang Song, Yunxin Zhao, University of Missouri, United States; Shaojun Wang, Ping An Technology, China

#5: CONFIDENCE ESTIMATION AND DELETION PREDICTION USING204 BIDIRECTIONAL RECURRENT NEURAL NETWORKS Anton Ragni, Qiujia Li, Mark Gales, Yu Wang, University of Cambridge, United Kingdom

#7: LSTM LANGUAGE MODEL ADAPTATION WITH IMAGES AND TITLES FOR219 MULTIMEDIA AUTOMATIC SPEECH RECOGNITION Yasufumi Moriya, Gareth Jones, Dublin City University, Ireland
#8: FIRST-PASS TECHNIQUES FOR VERY LARGE VOCABULARY SPEECH
#9: ANALYZING DEEP CNN-BASED UTTERANCE EMBEDDINGS FOR ACOUSTIC235 MODEL ADAPTATION Joanna Rownicka, Peter Bell, Steve Renals, The University of Edinburgh, United Kingdom
#10: MULTI-BAND PROCESSING WITH GABOR FILTERS AND TIME DELAY NEURAL242 NETS FOR NOISE ROBUST SPEECH RECOGNITION György Kovács, MTA-SZTE Research Group on Artificial Intelligence, Hungary; László Tóth, University of Szeged, Hungary; Gábor Gosztolya, MTA-SZTE Research Group on Artificial Intelligence, Hungary
#11: A TEACHER-STUDENT LEARNING APPROACH FOR UNSUPERVISED DOMAIN250 ADAPTATION OF SEQUENCE-TRAINED ASR MODELS <i>Vimal Manohar, Pegah Ghahremani, Daniel Povey, Sanjeev Khudanpur, Johns Hopkins University, United</i> <i>States</i>
#12: DOMAIN ROBUST FEATURE EXTRACTION FOR RAPID LOW RESOURCE ASR258 DEVELOPMENT Siddharth Dalmia, Xinjian Li, Florian Metze, Alan W Black, Carnegie Mellon University, United States
VOICE CONVERSION AND TTS
#1: STARGAN-VC: NON-PARALLEL MANY-TO-MANY VOICE CONVERSION USING266 STAR GENERATIVE ADVERSARIAL NETWORKS <i>Hirokazu Kameoka, Takuhiro Kaneko, Kou Tanaka, Nobukatsu Hojo, NTT Corporation, Japan</i>
#2: RHYTHM-FLEXIBLE VOICE CONVERSION WITHOUT PARALLEL DATA USING274 CYCLE-GAN OVER PHONEME POSTERIORGRAM SEQUENCES Cheng-chieh Yeh, Po-chun Hsu, Ju-chieh Chou, Hung-yi Lee, Lin-shan Lee, National Taiwan University, Taiwan

Berrak Sisman, Mingyang Zhang, National University of Singapore, Singapore; Sakriani Sakti, Nara Institute of Science and Technology, Japan; Haizhou Li, National University of Singapore, Singapore; Satoshi Nakamura, Nara Institute of Science and Technology, Japan

#5: AN EVALUATION OF DEEP SPECTRAL MAPPINGS AND WAVENET VOCODER297 FOR VOICE CONVERSION

Patrick Lumban Tobing, Tomoki Hayashi, Yi-Chiao Wu, Kazuhiro Kobayashi, Tomoki Toda, Nagoya University, Japan

Takuma Okamoto, National Institute of Information and Communications Technology, Japan; Tomoki Toda, Nagoya University, Japan; Yoshinori Shiga, Hisashi Kawai, National Institute of Information and Communications Technology, Japan

Fang-Yu Kuo, Sandesh Aryal, Gilles Degottex, Sam Kang, Pierre Lanchantin, Iris Ouyang, ObEN, Inc., United States

#9: COMPREHENSIVE EVALUATION OF STATISTICAL SPEECH WAVEFORM325 SYNTHESIS

Thomas Merritt, Bartosz Putrycz, Adam Nadolski, Tianjun Ye, Daniel Korzekwa, Wiktor Dolecki, Thomas Drugman, Viacheslav Klimkov, Alexis Moinet, Andrew Breen, Rafal Kuklinski, Nikko Strom, Roberto Barra-Chicote, Amazon, United Kingdom

#10: EXAMPLAR-BASED SPEECH WAVEFORM GENERATION FOR TEXT-TO-SPEECH332 *Cassia Valentini-Botinhao, Oliver Watts, Felipe Espic, Simon King, University of Edinburgh, United Kingdom*

ASR III (END-TO-END)

Jia Cui, Chao Weng, Guangsen Wang, Jun Wang, Tencent AI Lab, United States; Peidong Wang, The Ohio State University, United States; Chengzhu Yu, Dan Su, Dong Yu, Tencent AI Lab, United States

#2: COMBINING END-TO-END AND ADVERSARIAL TRAINING FOR LOW-RESOURCE361 SPEECH RECOGNITION

Jennifer Drexler, James Glass, Massachusetts Institute of Technology, United States

Shubham Toshniwal, Toyota Technological Institute at Chicago, United States; Anjuli Kannan, Chung-Cheng Chiu, Yonghui Wu, Tara N. Sainath, Google, United States; Karen Livescu, Toyota Technological Institute at Chicago, United States

#4: ON-DEVICE END-TO-END SPEECH RECOGNITION WITH MULTI-STEP376 PARALLEL RNNS

Yoonho Boo, Jinhwan Park, Lukas Lee, Wonyong Sung, Seoul National University, Republic of Korea

#5: DOMAIN ADAPTATION OF END-TO-END SPEECH RECOGNITION IN
#6: END-TO-END SPEECH RECOGNITION WITH WORD-BASED RNN LANGUAGE
#7: ACOUSTIC-TO-WORD RECOGNITION WITH SEQUENCE-TO-SEQUENCE
#8: COMBINING DE-NOISING AUTO-ENCODER AND RECURRENT NEURAL405 NETWORKS IN END-TO-END AUTOMATIC SPEECH RECOGNITION FOR NOISE ROBUSTNESS <i>Tzu-Hsuan Ting, Chia-Ping Chen, National Sun Yat-sen University, Taiwan</i>
#9: IMPROVED KNOWLEDGE DISTILLATION FROM BI-DIRECTIONAL TO411 UNI-DIRECTIONAL LSTM CTC FOR END-TO-END SPEECH RECOGNITION <i>Gakuto Kurata, Kartik Audhkhasi, IBM Research, Japan</i>
#10: DEEP CONTEXT: END-TO-END CONTEXTUAL SPEECH RECOGNITION418 Golan Pundak, Tara N. Sainath, Rohit Prabhavalkar, Anjuli Kannan, Ding Zhao, Google, United States
#11: BACK-TRANSLATION-STYLE DATA AUGMENTATION FOR END-TO-END ASR
#12: DIALOG-CONTEXT AWARE END-TO-END SPEECH RECOGNITION434 Suyoun Kim, Florian Metze, Carnegie Mellon University, United States
ASR IV
#1: TOWARD DOMAIN-INVARIANT SPEECH RECOGNITION VIA LARGE SCALE

TRAINING

Arun Narayanan, Ananya Misra, Khe Chai Sim, Golan Pundak, Anshuman Tripathi, Mohamed Elfeky, Parisa Haghani, Trevor Strohman, Michiel Bacchiani, Google, United States

#2: TRANSLITERATION BASED APPROACHES TO IMPROVE CODE-SWITCHED448 SPEECH RECOGNITION PERFORMANCE Jesse Emond, Bhuvana Ramabhadran, Brian Roark, Pedro Moreno, Min Ma, Google, United States

#3: EXPLORING LAYER TRAJECTORY LSTM WITH DEPTH PROCESSING UNITS456 AND ATTENTION

Jinyu Li, Liang Lu, Changliang Liu, Yifan Gong, Microsoft, United States

#4: MULTICHANNEL ASR WITH KNOWLEDGE DISTILLATION AND GENERALIZED463 CROSS CORRELATION FEATURE Wenjie Li, Yu Zhang, Pengyuan Zhang, Fengpei Ge, Key Laboratory of Speech Acoustics and Content Understanding, Institute of Acoustics, China
#5: OPTIMIZING THE QUALITY OF SYNTHETICALLY GENERATED470 PSEUDOWORDS FOR THE TASK OF MINIMAL-PAIR DISTINCTION Heiko Holz, Maria Chinkina, University of Tübingen, Germany; Laura Vetter, Ludwig Maximilian University of Munich, Germany
#6: LEVERAGING SEQUENCE-TO-SEQUENCE SPEECH SYNTHESIS FOR
#7: HIERARCHICAL MULTITASK LEARNING WITH CTC485 Ramon Sanabria, Florian Metze, Carnegie Mellon University, United States
#8: A K-NEAREST NEIGHBOURS APPROACH TO UNSUPERVISED SPOKEN TERM491 DISCOVERY Alaris Thual Corentin Dancette, Julian Karadavi, Juan Benjumaa, Emmanuel Dunoux, ENS, Erance
#9: A NEW TIMIT BENCHMARK FOR CONTEXT-INDEPENDENT PHONE
#10: EFFICIENT IMPLEMENTATION OF RECURRENT NEURAL NETWORK506 TRANSDUCER IN TENSORFLOW <i>Tom Bagby, Kanishka Rao, Khe Chai Sim, Google, United States</i>
#11: AUDIO-VISUAL SPEECH RECOGNITION WITH A HYBRID CTC/ATTENTION513 ARCHITECTURE Stavros Petridis, Imperial College London, United Kingdom; Themos Stafylakis, University of Nottingham, United Kingdom; Pingchuan Ma, Imperial College London, United Kingdom; Georgios Tzimiropoulos, University of Nottingham, United Kingdom; Maja Pantic, Imperial College London, United Kingdom
#12: MULTILINGUAL SEQUENCE-TO-SEQUENCE SPEECH RECOGNITION:
#13: SPEAKER SELECTIVE BEAMFORMER WITH KEYWORD MASK ESTIMATION528 <i>Yusuke Kida, Dung Tran, Motoi Omachi, Toru Taniguchi, Yuya Fujita, Yahoo Japan Corporation, Japan</i>
#14: SPEAKER ADAPTED BEAMFORMING FOR MULTI-CHANNEL AUTOMATIC535 SPEECH RECOGNITION Tobias Menne, Ralf Schlüter, Hermann Ney, RWTH Aachen University, Germany
#15: SPEAKER ADAPTATION FOR END-TO-END CTC MODELS

Peter Plantinga, Deblin Bagchi, Eric Fosler-Lussier, The Ohio State University, United States

Zhuo Chen, Xiong Xiao, Takuya Yoshioka, Jinyu Li, Hakan Erdogan, Yifan Gong, Microsoft Cloud & AI, United States

#18: A STUDY ON SPEECH ENHANCEMENT USING EXPONENT-ONLY FLOATING566 POINT QUANTIZED NEURAL NETWORK (EOFP-QNN)

Yi-Te Hsu, Academia Sinica, Taiwan; Yu-Chen Lin, Szu-Wei Fu, National Taiwan University, Taiwan; Yu Tsao, Academia Sinica, Taiwan; Tei-Wei Kuo, National Taiwan University, Taiwan

#19: RAPID SPEAKER ADAPTATION OF NEURAL NETWORK BASED FILTERBANK574 LAYER FOR AUTOMATIC SPEECH RECOGNITION

Hiroshi Seki, Toyohashi University of Technology, Japan; Kazumasa Yamamoto, Chubu University, Japan; Tomoyosi Akiba, Toyohashi University of Technology, Japan; Seiichi Nakagawa, Chubu University, Japan

#20: FAR-FIELD ASR USING LOW-RANK AND SPARSE SOFT TARGETS FROM581 PARALLEL DATA

Pranay Dighe, Idiap Research Institute, EPFL, Switzerland; Afsaneh Asaei, Idiap Research Institute, Switzerland; Herve Bourlard, Idiap Research Institute, EPFL, Switzerland

DEEP LEARNING FOR SPEECH SYNTHESIS

#2: A SPECTRALLY WEIGHTED MIXTURE OF LEAST SQUARE ERROR AND603 WASSERSTEIN DISCRIMINATOR LOSS FOR GENERATIVE SPSS Gilles Degottex, ObEN, Inc. - University of Cambridge, United Kingdom; Mark Gales, University of Cambridge, United Kingdom

#3: SCALING AND BIAS CODES FOR MODELING SPEAKER-ADAPTIVE DNN-BASED610 SPEECH SYNTHESIS SYSTEMS *Hieu-Thi Luong, Junichi Yamagishi, National Institute of Informatics, Japan*

#4: HIERARCHICAL RNNS FOR WAVEFORM-LEVEL SPEECH SYNTHESIS618 *Qingyun Dou, Moquan Wan, Gilles Degottex, Zhiyi Ma, Mark Gales, University of Cambridge, United Kingdom*

#5: PARAMETER GENERATION ALGORITHMS FOR TEXT-TO-SPEECH SYNTHESIS626 WITH RECURRENT NEURAL NETWORKS Viacheslav Klimkov, Alexis Moinet, Adam Nadolski, Thomas Drugman, Amazon, Poland

Kou Tanaka, Takuhiro Kaneko, Nobukatsu Hojo, Hirokazu Kameoka, NTT corporation, Japan

#7: IMPROVING UNSUPERVISED STYLE TRANSFER IN END-TO-END SPEECH640

SYNTHESIS WITH END-TO-END SPEECH RECOGNITION Da-Rong Liu, Chi-Yu Yang, Szu-Lin Wu, Hung-Yi Lee, National Taiwan University, Taiwan

#8: MULTI-SCALE ALIGNMENT AND CONTEXTUAL HISTORY FOR ATTENTION648

MECHANISM IN SEQUENCE-TO-SEQUENCE MODEL Andros Tjandra, Sakriani Sakti, Satoshi Nakamura, Nara Institute of Science and Technology, Japan

SPOKEN LANGUAGE UNDERSTANDING

Chunxi Liu, Matthew Wiesner, Shinji Watanabe, Craig Harman, Jan Trmal, Najim Dehak, Sanjeev Khudanpur, Johns Hopkins University, United States

Issac Alphonso, Nick Kibre, Tasos Anastasakos, Microsoft, United States

#3: A RE-RANKER SCHEME FOR INTEGRATING LARGE SCALE NLU MODELS......670 *Chengwei Su, Rahul Gupta, Shankar Ananthakrishnan, Spyros Matsoukas, Amazon, United States*

#4: PARSING COORDINATION FOR SPOKEN LANGUAGE UNDERSTANDING677 Sanchit Agarwal, Rahul Goel, Tagyoung Chung, Abhishek Sethi, Arindam Mandal, Spyros Matsoukas, Amazon, United States

#5: QUATERNION CONVOLUTIONAL NEURAL NETWORKS FOR THEME685 IDENTIFICATION OF TELEPHONE CONVERSATIONS

Titouan Parcollet, Mohamed Morchid, Georges Linarès, Université d'Avignon et des pays du Vaucluse, France; Renato De Mori, McGill University, Canada

#6: END-TO-END NAMED ENTITY AND SEMANTIC CONCEPT EXTRACTION FROM692 SPEECH

Sahar Ghannay, Antoine Caubrière, Yannick Estève, Nathalie Camelin, Edwin Simonnet, Antoine Laurent, University of Le Mans, France; Emmanuel Morin, University of Nantes, France

#7: TOWARD MULTI-FEATURES EMPHASIS SPEECH TRANSLATION: ASSESSMENT700 OF HUMAN EMPHASIS PRODUCTION AND PERCEPTION WITH SPEECH AND TEXT CLUES

Quoc Truong Do, Nara Institute of Science and Technology, Japan; Sakriani Sakti, Satoshi Nakamura, Nara Institute of Science and Technology/AIP, Japan

#8: CORPUS AND ANNOTATION TOWARDS NLU FOR CUSTOMER ORDERING707 DIALOGS

John Chen, Rashmi Prasad, Svetlana Stoyanchev, Ethan Selfridge, Srinivas Bangalore, Michael Johnston, Interactions, LLC., United States

#9: COUPLED REPRESENTATION LEARNING FOR DOMAINS, INTENTS AND SLOTS714 IN SPOKEN LANGUAGE UNDERSTANDING

Jihwan Lee, Dongchan Kim, Ruhi Sarikaya, Young-Bum Kim, Amazon, United States

#10: FROM AUDIO TO SEMANTICS: APPROACHES TO END-TO-END SPOKEN720 LANGUAGE UNDERSTANDING Parisa Haghani, Arun Narayanan, Michiel Bacchiani, Galen Chuang, Neeraj Gaur, Pedro Moreno, Rohit Prabhavalkar, Zhongdi Qu, Austin Waters, Google, United States #11: INVESTIGATING THE DOWNSTREAM IMPACT OF GRAPHEME-BASED727 ACOUSTIC MODELING ON SPOKEN UTTERANCE CLASSIFICATION Rvan Price, Bhargav Srinivas Ch, Surbhi Singhal, Srinivas Bangalore, Interactions, LLC., United States NATURAL LANGUAGE PROCESSING #1: ABSTRACTIVE DIALOGUE SUMMARIZATION WITH SENTENCE-GATED735 MODELING OPTIMIZED BY DIALOGUE ACTS Chih-Wen Goo, Yun-Nung Chen, National Taiwan University, Taiwan #2: GRAPH-BASED DEEP-TREE RECURSIVE NEURAL NETWORK (DTRNN) FOR743 TEXT CLASSIFICATION Fenxiao Chen, Bin Wang, C.-C. Jay Kuo, University of Southern California, United States #3: EXTENSION OF CONVENTIONAL CO-TRAINING LEARNING STRATEGIES TO750 THREE-VIEW AND COMMITTEE-BASED LEARNING STRATEGIES FOR EFFECTIVE AUTOMATIC SENTENCE SEGMENTATION Dogan Dalva, Umit Guz, F.M.V. ISIK University, Turkey; Hakan Gurkan, Bursa Technical University, Turkey Lyan Verwimp, KU Leuven, Belgium; Joris Pelemans, Apple, United States; Hugo Van hamme, Patrick Wambacq, KU Leuven, Belgium #5: WORD SEGMENTATION FROM PHONEME SEQUENCES BASED ON763 PITMAN-YOR SEMI-MARKOV MODEL EXPLOITING SUBWORD INFORMATION Ryu Takeda, Kazunori Komatani, Osaka University, Japan; Alexander Rudnicky, Carnegie Mellon University, United States **QUESTION GENERATION** Lasguido Nio, Koji Murakami, Rakuten Institute of Technology, Japan **#7: INVESTIGATING LINGUISTIC PATTERN ORDERING IN HIERARCHICAL779** NATURAL LANGUAGE GENERATION Shang-Yu Su, Yun-Nung Chen, National Taiwan University, Taiwan OF AUTHOR-STYLIZED POETRY Alexey Tikhonov, Yandex, Germany; Ivan Yamshchikov, Max Planck Institute for Mathematics in the Sciences, Germany **#9: GENERATING SEMANTIC SIMILARITY ATLAS FOR NATURAL LANGUAGES795** Lütfi Kerem Şenel, İhsan Utlu, Veysel Yücesoy, Aykut Koç, ASELSAN, Turkey; Tolga Çukur, Bilkent University, Turkey ELEMENTS

Nevan Wichers, Dilek Hakkani-Tür, Jindong (JD) Chen, Google AI, United States

#11: SENTIMENT CLASSIFICATION ON ERRONEOUS ASR TRANSCRIPTS: A MULTI807 VIEW LEARNING APPROACH

Sri Harsha Dumpala, Imran Sheikh, Rupayan Chakraborty, Sunil Kumar Kopparapu, TCS Research and Innovation-Mumbai, India

DIALOGUE

#5: CONTEXT-AWARE DIALOG RE-RANKING FOR TASK-ORIENTED DIALOG846 SYSTEMS

Junki Ohmura, Sony Corporation, United States; Maxine Eskenazi, Language Technologies Institute Carnegie Mellon University, United States

Yuka Kobayashi, Takami Yoshida, Kenji Iwata, Hiroshi Fujimura, Masami Akamine, Toshiba Corporation, Japan

Rui Zhao, Volker Tresp, Siemens & LMU, Germany

#9: OPTIMIZING NEURAL RESPONSE GENERATOR WITH EMOTIONAL IMPACT876 INFORMATION

Nurul Lubis, Sakriani Sakti, Koichiro Yoshino, Satoshi Nakamura, Nara Institute of Science and Technology, Japan

Chandra Khatri, Rahul Goel, Behnam Hedayatnia, Angeliki Metanillou, Anushree Venkatesh, Raefer Gabriel, Arindam Mandal, Amazon Alexa, United States

Izzeddin Gur, University of California Santa Barbara, United States; Dilek Hakkani-Tür, Google, United States; Gokhan Tür, Uber AI Labs, United States; Pararth Shah, Facebook, United States

#13: SCALABLE LANGUAGE MODEL ADAPTATION FOR SPOKEN DIALOGUE SYSTEMS907 *Ankur Gandhe, Ariya Rastrow, Björn Hoffmeister, Amazon, United States*

CORPORA AND EVALUATION METHODOLOGIES

#1: DISCOURSE MODELING OF NON-NATIVE SPONTANEOUS SPEECH USING913 THE RHETORICAL STRUCTURE THEORY FRAMEWORK

Xinhao Wang, Binod Gyawali, James V. Bruno, Hillary R. Molloy, Keelan Evanini, Klaus Zechner, Educational Testing Service, United States

#4: INVESTIGATION OF USERS' SHORT RESPONSES IN ACTUAL CONVERSATION934 SYSTEM AND AUTOMATIC RECOGNITION OF THEIR INTENTIONS

Katsuya Yokoyama, Hiroaki Takatsu, Waseda University, Japan; Hiroshi Honda, Honda R&D Co., Ltd, Japan; Shinya Fujie, Chiba Institute of Technology, Japan; Tetsunori Kobayashi, Waseda University, Japan

#5: PHONETIC-AND-SEMANTIC EMBEDDING OF SPOKEN WORDS WITH941 APPLICATIONS IN SPOKEN CONTENT RETRIEVAL

Yi-Chen Chen, Sung-Feng Huang, Chia-Hao Shen, Hung-yi Lee, Lin-shan Lee, National Taiwan University, Taiwan

Marco Del Vecchio, Andrey Malinin, Mark Gales, University of Cambridge, United Kingdom

#8: EVALUATING ON-DEVICE ASR ON FIELD RECORDINGS FROM AN964 INTERACTIVE READING COMPANION

Anastassia Loukina, Nitin Madnani, Beata Beigman Klebanov, Abhinav Misra, Educational Testing Service, United States; Georgi Angelov, Astea Solutions, Bulgaria; Ognjen Todic, Keen Research, United States

Yao Qian, Rutuja Ubale, Matthew Mulholland, Keelan Evanini, Xinhao Wang, Educational Testing Service, United States

Joana Correia, Carnegie Mellon University / INESC-ID, Portugal; Isabel Trancoso, INESC-ID / IST, Portugal; Bhiksha Raj, Carnegie Mellon University, Portugal

SPEAKER RECOGNITION/VERIFICATION

#1: TEXT-INDEPENDENT SPEAKER VERIFICATION BASED ON DEEP NEURAL1001 NETWORKS AND SEGMENTAL DYNAMIC TIME WARPING

Mohamed Adel, Mohamed Afify, Microsoft Advanced Technology Lab, Cairo, Egypt; Akram Gaballah, Microsoft Corporation, United States; Magda Fayek, Cairo University, Egypt

Suwon Shon, Hao Tang, James Glass, Massachusetts Institute of Technology, United States

#3: GENERATIVE X-VECTORS FOR TEXT-INDEPENDENT SPEAKER VERIFICATION1014 Longting Xu, Rohan Kumar Das, Emre Yilmaz, Jichen Yang, Haizhou Li, National University of Singapore, Singapore

#4: SPEAKER RECOGNITION FROM RAW WAVEFORM WITH SINCNET......**1021** *Mirco Ravanelli, Yoshua Bengio, Université de Montréal, Canada*

#5: INVESTIGATING DEEP NEURAL NETWORKS FOR SPEAKER DIARIZATION IN1029 THE DIHARD CHALLENGE

Ivan Himawan, Md Hafizur Rahman, Sridha Sridharan, Clinton Fookes, Ahilan Kanagasundaram, Queensland University of Technology, Australia

#6: ROLE ANNOTATED SPEECH RECOGNITION FOR CONVERSATIONAL1036 INTERACTIONS

Nikolaos Flemotomos, Zhuohao Chen, University of Southern California, United States; David Atkins, University of Washington, United States; Shrikanth Narayanan, University of Southern California, United States

#7: TEACHER-STUDENT TRAINING FOR TEXT-INDEPENDENT SPEAKER1044 RECOGNITION

Raymond W. M. Ng, Xuechen Liu, Emotech Labs, United Kingdom; Pawel Swietojanski, The University of New South Wales, Australia

#8: ATTENTION MECHANISM IN SPEAKER RECOGNITION: WHAT DOES IT LEARN1052 IN DEEP SPEAKER EMBEDDING?

Qiongqiong Wang, Koji Okabe, Kong Aik Lee, Hitoshi Yamamoto, Takafumi Koshinaka, NEC Corporation, Japan

#9: DETECTION AND CALIBRATION OF WHISPER FOR SPEAKER RECOGNITION1060 *Finnian Kelly, John H.L. Hansen, The University of Texas at Dallas, United States*

#10: TRAINING SPEAKER RECOGNITION MODELS WITH RECORDING-LEVEL1066 LABELS

Tanel Alumäe, Tallinn University of Technology, Estonia

#11: SHORT UTTERANCE SPEAKER RECOGNITION BY RESERVOIR WITH1073 SELF-ORGANIZED MAPPING

Narumitsu Ikeda, The University of Tokyo, Japan; Yoshinao Sato, Fairy Devices Inc., Japan; Hirokazu Takahashi, The University of Tokyo, Japan