2019 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU 2019)

Singapore
14-18 December 2019

Pages 1-538

IEEE Catalog Number: CFP19SRW-POD
ISBN: 978-1-7281-0307-5

Copyright © 2019 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: CFP19SRW-POD
ISBN (Print-On-Demand): 978-1-7281-0307-5
ISBN (Online): 978-1-7281-0306-8

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

ADV: ASR IN ADVERSE ENVIRONMENTS

ADV.1: SPEAKER-AWARE SPEECH-TRANSFORMER.....222
Zhiyun Fan, Institute of Automation, Chinese Academy of Sciences, China; Jie Li, Kwai, China; Shiyu Zhou, Bo Xu, Institute of Automation, Chinese Academy of Sciences, China

ADV.2: SPEECH SEPARATION USING SPEAKER INVENTORY.....230
Peidong Wang, Ohio State University, United States; Zhuo Chen, Xiong Xiao, Zhong Meng, Takuya Yoshioka, Tianyan Zhou, Liang Lu, Jinyu Li, Microsoft, United States

ADV.3: MIMO-SPEECH: END-TO-END MULTI-CHANNEL MULTI-SPEAKER SPEECH RECOGNITION 237
Xuankai Chang, Johns Hopkins University, United States; Wangyou Zhang, Yanmin Qian, Shanghai Jiao Tong University, China; Jonathan Le Roux, Mitsubishi Electric Research Laboratories (MERL), United States; Shinji Watanabe, Johns Hopkins University, United States

ADV.4: JOINT DISTRIBUTION LEARNING IN THE FRAMEWORK OF VARIATIONAL AUTOENCODERS FOR FAR-FIELD SPEECH ENHANCEMENT 245
Mahesh Kumar Chelimilla, Shashi Kumar, Shakti P. Rath, Samsung Research Institute Bangalore, India

ADV.5: ANALYZING LARGE RECEPTIVE FIELD CONVOLUTIONAL NETWORKS FOR DISTANT SPEECH RECOGNITION 252
Salar Jafarlou Jafarlou, Soheil Khorram, Vinay Kothapally, John H.L. Hansen, University of Texas at Dallas, United States

ADV.6: FASNET: LOW-LATENCY ADAPTIVE BEAMFORMING FOR MULTI-MICROPHONE AUDIO PROCESSING 260
Yi Luo, Columbia University, United States; Enea Ceolini, University of Zurich and ETH Zurich, Switzerland; Cong Han, Columbia University, United States; Shih-Chii Liu, University of Zurich and ETH Zurich, Switzerland; Nima Mesgarani, Columbia University, United States

ADV.7: DOMAIN ADAPTATION VIA TEACHER-STUDENT LEARNING FOR END-TO-END SPEECH RECOGNITION 268
Zhong Meng, Jinyu Li, Yashesh Gaur, Yifan Gong, Microsoft Corporation, United States

ADV.8: ADVANCES IN ONLINE AUDIO-VISUAL MEETING TRANSCRIPTION 276
Takuya Yoshioka, Igor Abramovski, Cem Aksoylar, Zhuo Chen, Moshe David, Dimitrios Dimitriadis, Yifan Gong, Ilya Gurvich, Xuedong Huang, Yan Huang, Aviv Hurvitz, Li Jiang, Sharon Koubi, Eyal Krupka, Ido Leichter, Changliang Liu, Partha Parthasarathy, Alon Vinnikov, Lingfeng Wu, Xiong Xiao, Wayne Xiong, Huaming Wang, Zhenghao Wang, Jun Zhang, Yong Zhao, Tianyan Zhou, Microsoft, United States

ASR-1: AUTOMATIC SPEECH RECOGNITION I

ASR-1.1: INCREMENTAL LATTICE DETERMINIZATION FOR WFST DECODERS 1
Zhehuai Chen, Shanghai Jiao Tong University, China; Mahsa Yarmohammadi, Hainan Xu, Johns Hopkins University, United States; Hang Lv, Lei Xie, Northwestern Polytechnical University, China; Daniel Povey, Sanjeev Khudanpur, Johns Hopkins University, United States

ASR-1.2: A COMPARISON OF TRANSFORMER AND LSTM ENCODER DECODER MODELS FOR ASR 8
Albert Zeyer, Parnia Bahar, Kazuki Irie, Ralf Schlüter, Hermann Ney, RWTH Aachen University, Germany

ASR-1.3: A DROPOUT-BASED SINGLE MODEL COMMITTEE APPROACH FOR ACTIVE LEARNING IN ASR 16
Jiayi Fu, Kuang Ru, Zhuiyi Technology Company, China

ASR-1.4: PERSONALIZATION OF END-TO-END SPEECH RECOGNITION ON MOBILE DEVICES FOR NAMED ENTITIES 23
Khe Chai Sim, Francoise Beaufays, Arnaud Benard, Dhruv Guliani, Andreas Kabel, Nikhil Khare, Tamar Lucassen, Petr Zadrazil, Harry Zhang, Leif Johnson, Giovanni Motta, Lillian Zhou, Google, United States

ASR-1.5: SIMULTANEOUS SPEECH RECOGNITION AND SPEAKER DIARIZATION FOR MONAURAL DIALOGUE RECORDINGS WITH TARGET-SPEAKER ACOUSTIC MODELS 31
Naoyuki Kanda, Shota Horiguchi, Yusuke Fujita, Yawen Xue, Kenji Nagamatsu, Hitachi, Ltd., Japan; Shinji Watanabe, Johns Hopkins University, United States

ASR-1.6: INTEGRATING SOURCE-CHANNEL AND ATTENTION-BASED SEQUENCE-TO-SEQUENCE MODELS FOR SPEECH RECOGNITION 39
Qiujia Li, Chao Zhang, Phil Woodland, University of Cambridge, United Kingdom

ASR-1.7: AN INVESTIGATION INTO THE EFFECTIVENESS OF ENHANCEMENT IN ASR TRAINING AND TEST FOR CHIME-5 DINNER PARTY TRANSCRIPTION 47
Catalin Zorila, Toshiba Cambridge Research Laboratory, United Kingdom; Christoph Boeddeker, Paderborn University, Germany; Rama Doddipatla, Toshiba Cambridge Research Laboratory, United Kingdom; Reinhold Haeb-Umbach, Paderborn University, Germany

ASR-1.8: STATE-OF-THE-ART SPEECH RECOGNITION USING MULTI-STREAM SELF-ATTENTION WITH DILATED 1D CONVOLUTIONS 54
Kyu Han, Ramon Prieto, Tao Ma, ASAPP, Inc., United States

ASR-1.9: HIGHLY EFFICIENT NEURAL NETWORK LANGUAGE MODEL COMPRESSION USING SOFT BINARIZATION TRAINING 62
Rao Ma, Qi Liu, Kai Yu, Shanghai Jiao Tong University, China

ASR-1.10: IMPROVED MULTI-STAGE TRAINING OF ONLINE ATTENTION-BASED ENCODER-DECODER MODELS 70
Abhinav Garg, Dhananjaya Gowda, Ankur Kumar, Kwangyoun Kim, Mehul Kumar, Chanwoo Kim, Samsung Research, Korea (South)

ASR-1.11: LEAD2GOLD: TOWARDS EXPLOITING THE FULL POTENTIAL OF NOISY TRANSCRIPTIONS FOR SPEECH RECOGNITION 78
Adrien Dufraux, Facebook AI Research, France; Emmanuel Vincent, INRIA, France; Awni Hannun, Facebook AI Research, United States; Armelle Brun, Université de Lorraine, France; Matthijs Douze, Facebook AI Research, France

ASR-1.12: ORTHOGONALITY CONSTRAINED MULTI-HEAD ATTENTION FOR KEYWORD SPOTTING 86
Mingu Lee, Jinkyu Lee, Hye Jin Jang, Byeonggeun Kim, Wonil Chang, Kyuwoong Hwang, Qualcomm AI Research, Korea (South)

ASR-1.13: LEARNING BETWEEN DIFFERENT TEACHER AND STUDENT MODELS IN ASR 93
Jeremy Heng Meng Wong, Microsoft, United States; Mark John Francis Gales, Yu Wang, University of Cambridge, United Kingdom

ASR-1.14: A UNIFIED ENDPOINTER USING MULTITASK AND MULTIDOMAIN TRAINING 100
Shuo-Yiin Chang, Bo Li, Gabor Simko, Google, United States

ASR-1.15: DOMAIN EXPANSION IN DNN-BASED ACOUSTIC MODELS FOR ROBUST SPEECH RECOGNITION 107
Shahram Ghorbani, Soheil Khorram, John H.L. Hansen, University of Texas at Dallas, United States

ASR-1.16: IMPROVING RNN TRANSDUCER MODELING FOR END-TO-END SPEECH RECOGNITION 114
Jinyu Li, Rui Zhao, Hu Hu, Yifan Gong, Microsoft, United States

ASR-1.17: SIMPLE GATED CONVNET FOR SMALL FOOTPRINT ACOUSTIC MODELING 122
Lukas Lee, Jinhwan Park, Wonyong Sung, Seoul National University, Korea (South)

ASR-1.18: GANS FOR CHILDREN: A GENERATIVE DATA AUGMENTATION STRATEGY FOR CHILDREN SPEECH RECOGNITION 129
Peiyao Sheng, Zhuolin Yang, Yanmin Qian, Shanghai Jiao Tong University, China

ASR-1.19: ESPRESSO: A FAST END-TO-END NEURAL SPEECH RECOGNITION TOOLKIT 136
Yiming Wang, Tongfei Chen, Hainan Xu, Shuoyang Ding, Johns Hopkins University, United States; Hang Lv, Northwestern Polytechnical University, China; Yiwen Shao, Johns Hopkins University, United States; Nanyun Peng, University of Southern California, United States; Lei Xie, Northwestern Polytechnical University, China; Shinji Watanabe, Sanjeev Khudanpur, Johns Hopkins University, United States

ASR-2: AUTOMATIC SPEECH RECOGNITION II

ASR-2.1: TRAINING LANGUAGE MODELS FOR LONG-SPAN CROSS-SENTENCE EVALUATION 419
Kazuki Irie, Albert Zeyer, Ralf Schlüter, Hermann Ney, RWTH Aachen University, Germany

ASR-2.2: TRANSFORMER ASR WITH CONTEXTUAL BLOCK PROCESSING 427
Emiru Tsunoo, Yosuke Kashiwagi, Toshiyuki Kumakura, Sony Corporation, Japan; Shinji Watanabe, Johns Hopkins University, United States

ASR-2.3: A DENSITY RATIO APPROACH TO LANGUAGE MODEL FUSION IN END-TO-END AUTOMATIC SPEECH RECOGNITION 434
Erik McDermott, Hasim Sak, Ehsan Variani, Google Inc, United States

ASR-2.4: IMPROVING GRAPHEME-TO-PHONEME CONVERSION BY INVESTIGATING COPYING MECHANISM IN RECURRENT ARCHITECTURES 442
Abhishek Niranjan, Mahaboob Ali Basha Shaik, Samsung Research and Development Institute, India

ASR-2.5: A COMPARATIVE STUDY ON TRANSFORMER VS RNN IN SPEECH APPLICATIONS 449
Shigeki Karita, NTT Communication Science Laboratories, Japan; Nanxin Chen, Johns Hopkins University, United States; Tomoki Hayashi, Nagoya University, Japan; Takaaki Hori, Mitsubishi Electric Research Laboratories (MERL), United States; Hirofumi Inaguma, Kyoto University, Japan; Ziyan Jiang, Johns Hopkins University, United States; Masao Someki, Nagoya University, Japan; Nelson Enrique Yalta Soplin, Waseda University, Japan; Ryuichi Yamamoto, LINE Corporation, Japan; Xiaofei Wang, Shinji Watanabe, Johns Hopkins University, United States; Takenori Yoshimura, Nagoya University, Japan; Wangyou Zhang, Shanghai Jiao Tong University, China

ASR-2.6: FROM SENONES TO CHENONES: TIED CONTEXT-DEPENDENT GRAPHEMES FOR HYBRID SPEECH RECOGNITION 457
Duc Le, Xiaohui Zhang, Weiyi Zheng, Christian Fuegen, Geoffrey Zweig, Michael L. Seltzer, Facebook, United States

ASR-2.7: ATTENTION-BASED SPEECH RECOGNITION USING GAZE INFORMATION 465
Osamu Segawa, Chubu Electric Power Co., Inc., Japan; Tomoki Hayashi, Kazuya Takeda, Nagoya University, Japan

ASR-2.8: LISTENING WHILE SPEAKING AND VISUALIZING: IMPROVING ASR THROUGH MULTIMODAL CHAIN 471
Johanes Effendi, Nara Institute of Science and Technology / RIKEN Center for Advanced Intelligence Project AIP, Japan; Andros Tjandra, Nara Institute of Science and Technology, Japan; Sakriani Sakti, Satoshi Nakamura, Nara Institute of Science and Technology / RIKEN Center for Advanced Intelligence Project AIP, Japan

ASR-2.9: EMBEDDINGS FOR DNN SPEAKER ADAPTIVE TRAINING 479
Joanna Rownicka, Peter Bell, Steve Renals, University of Edinburgh, United Kingdom

ASR-2.10: LANGUAGE MODEL BOOTSTRAPPING USING NEURAL MACHINE TRANSLATION FOR CONVERSATIONAL SPEECH RECOGNITION 487
Surabhi Punjabi, Harish Arsikere, Sri Garimella, Amazon, India

ASR-2.11: SPEAKER AND LANGUAGE AWARE TRAINING FOR END-TO-END ASR 494
Shubham Bansal, Karan Malhotra, Sriram Ganapathy, Indian Institute of Science, Bangalore, India

ASR-2.12: DATA AUGMENTATION BASED ON VOWEL STRETCH FOR IMPROVING CHILDREN'S SPEECH RECOGNITION 502
Tohru Nagano, Takashi Fukuda, Masayuki Suzuki, Gakuto Kurata, IBM, Japan

ASR-2.13: MIXED BANDWIDTH ACOUSTIC MODELING LEVERAGING KNOWLEDGE DISTILLATION 509
Takashi Fukuda, Samuel Thomas, IBM, Japan

ASR-2.14: ON TEMPORAL CONTEXT INFORMATION FOR HYBRID BLSTM-BASED PHONEME RECOGNITION 516
Timo Lohrenz, Maximilian Strake, Tim Fingscheidt, Technische Universität Braunschweig, Germany

ASR-2.15: EXPLORING MODEL UNITS AND TRAINING STRATEGIES FOR END-TO-END SPEECH RECOGNITION 524
Mingkun Huang, YiZhou Lu, Shanghai Jiao Tong University, China; Lan Wang, Chinese Academy of Sciences, China; Yanmin Qian, Kai Yu, Shanghai Jiao Tong University, China

ASR-2.16: QUERY-BY-EXAMPLE ON-DEVICE KEYWORD SPOTTING.....532
Byeonggeun Kim, Mingu Lee, Jinkyu Lee, Yeonseok Kim, Kyuwoong Hwang, Qualcomm, Korea (South)

ASR-2.17: SMALL-FOOTPRINT KEYWORD SPOTTING WITH GRAPH CONVOLUTIONAL NETWORK 539
Xi Chen, Shouyi Yin, Tsinghua University, China; Dandan Song, Peng Ouyang, TsingMicro Co. Ltd., China; Leibo Liu, Shaojun Wei, Tsinghua University, China

ASR-2.18: SIMPLIFIED LSTMS FOR SPEECH RECOGNITION.....547
George Saon, Zoltan Tuske, Kartik Audhkhasi, Brian Kingsbury, Michael Picheny, Samuel Thomas, IBM, United States

ASR-2.19: GENERALIZED LARGE-CONTEXT LANGUAGE MODELS BASED ON FORWARD-BACKWARD HIERARCHICAL RECURRENT ENCODER-DECODER MODELS 554
Ryo Masumura, Mana Ihori, Tomohiro Tanaka, Itsumi Saito, Kyosuke Nishida, Takanobu Oba, NTT Corporation, Japan

ASR-2.20: END-TO-END TRAINING OF A LARGE VOCABULARY END-TO-END SPEECH RECOGNITION SYSTEM 562
Chanwoo Kim, Sungsoo Kim, Kwangyoun Kim, Mehul Kumar, Jiyeon Kim, Kyungmin Lee, Changwoo Han, Abhinav Garg, Eunhyang Kim, Minkyoo Shin, Shatrughan Singh, Larry Heck, Dhananjaya Gowda, Samsung Research, Korea (South)

ASR-3: AUTOMATIC SPEECH RECOGNITION III

ASR-3.1: SEMI-SUPERVISED TRAINING AND DATA AUGMENTATION FOR ADAPTATION OF AUTOMATIC BROADCAST NEWS CAPTIONING SYSTEMS 867
Yinghui Huang, Samuel Thomas, Masayuki Suzuki, Zoltan Tuske, Larry Sansone, Michael Picheny, IBM, United States

ASR-3.2: ONLINE BATCH NORMALIZATION ADAPTATION FOR AUTOMATIC SPEECH RECOGNITION 875
Franco Mana, Felix Weninger, Roberto Gemello, Puming Zhan, NUANCE Communications, Italy

ASR-3.3: SPEAKER ADAPTIVE TRAINING USING MODEL AGNOSTIC META-LEARNING.....881
Ondrej Klejch, Joachim Fainberg, Peter Bell, Steve Renals, University of Edinburgh, United Kingdom

ASR-3.4: A COMPARISON OF END-TO-END MODELS FOR LONG-FORM SPEECH RECOGNITION 889
Chung-Cheng Chiu, Wei Han, Yu Zhang, Ruoming Pang, Sergey Kishchenko, Google Inc, United States; Patrick Nguyen, Grab Technologies, United States; Hagen Soltau, Arun Narayanan, Hank Liao, Shuyuan Zhang, Anjuli Kannan, Rohit Prabhavalkar, Zhifeng Chen, Tara Sainath, Yonghui Wu, Google Inc, United States

ASR-3.5: ACOUSTIC MODEL ADAPTATION FROM RAW WAVEFORMS WITH SINCNET.....897
Joachim Fainberg, Ondrej Klejch, Erfan Loweimi, Peter Bell, Steve Renals, University of Edinburgh, United Kingdom

ASR-3.6: RECURRENT NEURAL NETWORK TRANSDUCER FOR AUDIO-VISUAL SPEECH RECOGNITION 905
Takaki Makino, Hank Liao, Google Inc., United States; Yannis Assael, Brendan Shillingford, DeepMind, United Kingdom; Basilio Garcia, Otavio Braga, Olivier Siohan, Google Inc., United States

ASR-3.7: EXPLICIT ALIGNMENT OF TEXT AND SPEECH ENCODINGS FOR ATTENTION-BASED END-TO-END SPEECH RECOGNITION 913
Jennifer Drexler, James Glass, Massachusetts Institute of Technology, United States

ASR-3.8: RECOGNIZING LONG-FORM SPEECH USING STREAMING END-TO-END MODELS 920
Arun Narayanan, Rohit Prabhavalkar, Chung-Cheng Chiu, David Rybach, Tara Sainath, Trevor Strohman, Google Inc., United States

ASR-3.9: LEVERAGING LANGUAGE ID IN MULTILINGUAL END-TO-END SPEECH RECOGNITION 928
Austin Waters, Neeraj Gaur, Parisa Haghani, Pedro Moreno, Zhongdi Qu, Google, United States

ASR-3.10: STREAMING END-TO-END SPEECH RECOGNITION WITH JOINT CTC-ATTENTION BASED MODELS 936
Niko Moritz, Takaaki Hori, Jonathan Le Roux, Mitsubishi Electric Research Laboratories (MERL), United States

ASR-3.11: MONOTONIC RECURRENT NEURAL NETWORK TRANSDUCER AND DECODING STRATEGIES 944
Anshuman Tripathi, Han Lu, Hasim Sak, Hagen Soltau, Google, United States

ASR-3.12: CHARACTER-AWARE ATTENTION-BASED END-TO-END SPEECH RECOGNITION 949
Zhong Meng, Yashesh Gaur, Jinyu Li, Yifan Gong, Microsoft Corporation, United States

ASR-3.13: ATTENTION BASED ON-DEVICE STREAMING SPEECH RECOGNITION WITH LARGE SPEECH CORPUS 956
Kwangyoun Kim, Kyungmin Lee, Dhananjaya Gowda, Junmo Park, Sungsoo Kim, Sichen Jin, Young-Yoon Lee, Jinsu Yeo, Daehyun Kim, Seokyeong Jung, Jungin Lee, Myoungji Han, Chanwoo Kim, Samsung Electronics, Korea (South)

ASR-3.14: ZERO-SHOT CODE-SWITCHING ASR AND TTS WITH MULTILINGUAL MACHINE SPEECH CHAIN 964
Sahoko Nakayama, Nara Institute of Science and Technology / RIKEN Center for Advanced Intelligence Project AIP, Japan; Andros Tjandra, Nara Institute of Science and Technology, Japan; Sakriani Sakti, Satoshi Nakamura, Nara Institute of Science and Technology / RIKEN Center for Advanced Intelligence Project AIP, Japan

ASR-3.15: END-TO-END CODE-SWITCHING ASR FOR LOW-RESOURCED LANGUAGE PAIRS 972
Xianghu Yue, Beijing Institute of Technology / National University of Singapore, Singapore; Grandee Lee, Emre Yilmaz, National University of Singapore, Singapore; Fang Deng, Beijing Institute of Technology, China; Haizhou Li, National University of Singapore, Singapore

ASR-3.16: UNSUPERVISED ADAPTATION OF ACOUSTIC MODELS FOR ASR USING UTTERANCE-LEVEL EMBEDDINGS FROM SQUEEZE AND EXCITATION NETWORKS 980
Hardik Sailor, Salil Deena, Md Asif Jalal, Rasa Lileikyte, Thomas Hain, University of Sheffield, United Kingdom

ASR-3.17: POWER-LAW NONLINEARITY WITH MAXIMALLY UNIFORM DISTRIBUTION CRITERION FOR IMPROVED NEURAL NETWORK TRAINING IN AUTOMATIC SPEECH RECOGNITION 988
Chanwoo Kim, Mehul Kumar, Kwangyoun Kim, Dhananjaya Gowda, Samsung Research, Korea (South)

ASR-3.18: SPEECH RECOGNITION WITH AUGMENTED SYNTHESIZED SPEECH 996
Andrew Rosenberg, Yu Zhang, Bhuvana Ramabhadran, Ye Jia, Pedro Moreno, Yonghui Wu, Zelin Wu, Google, United States

MLP: MULTILINGUAL LANGUAGE PROCESSING

MLP.1: SECOND LANGUAGE TRANSFER LEARNING IN HUMANS AND MACHINES USING IMAGE SUPERVISION 1040
Kiran Praveen, Indian Institute of Science, India; Anshul Gupta, International Institute of Information Technology, Bangalore, India; Akshara Soman, Sriram Ganapathy, Indian Institute of Science, India

MLP.2: ZERO-SHOT PRONUNCIATION LEXICONS FOR CROSS-LANGUAGE ACOUSTIC MODEL TRANSFER 1048
Matthew Wiesner, Oliver Adams, David Yarowsky, Jan Trmal, Sanjeev Khudanpur, Johns Hopkins University, United States

NEW: NEW APPLICATIONS OF ASR

NEW.1: IMPROVING SPEECH ENHANCEMENT WITH PHONETIC EMBEDDING FEATURES 645
Bo Wu, Meng Yu, Lianwu Chen, Mingjie Jin, Dan Su, Dong Yu, Tencent, China

NEW.2: DETECTING DECEPTION IN POLITICAL DEBATES USING ACOUSTIC AND TEXTUAL FEATURES 652
Daniel Kopev, Sofia University, Bulgaria; Ahmed Ali, Hamad Bin Khalifa University, Qatar; Ivan Koychev, Sofia University, Bulgaria; Preslav Nakov, Hamad Bin Khalifa University, Qatar

NEW.3: END-TO-END OVERLAPPED SPEECH DETECTION AND SPEAKER COUNTING WITH RAW WAVEFORM 660
Wangyou Zhang, Man Sun, Shanghai Jiao Tong University, China; Lan Wang, Chinese Academy of Sciences, China; Yanmin Qian, Shanghai Jiao Tong University, China

NEW.4: TIME DOMAIN AUDIO VISUAL SPEECH SEPARATION 667
Jian Wu, Northwestern Polytechnical University, China; Yong Xu, Shi-Xiong Zhang, Lianwu Chen, Meng Yu, Tencent AI Lab, United States; Lei Xie, Northwestern Polytechnical University, China; Dong Yu, Tencent AI Lab, United States

NEW.5: SPEECH REVEALS FUTURE RISK OF DEVELOPING DEMENTIA: PREDICTIVE DEMENTIA SCREENING FROM BIOGRAPHIC INTERVIEWS 674
Jochen Weiner, University of Bremen, Germany; Claudia Frankenberg, Johannes Schröder, University of Heidelberg, Germany; Tanja Schultz, University of Bremen, Germany

NEW.6: IMPROVING FUNDAMENTAL FREQUENCY GENERATION IN EMG-TO-SPEECH CONVERSION USING A QUANTIZATION APPROACH 682
Lorenz Diener, Tejas Umesh, Tanja Schultz, University of Bremen, Germany

NEW.7: TOWARDS REAL-TIME MISPRONUNCIATION DETECTION IN KIDS' SPEECH 690
Peter Plantinga, Eric Fosler-Lussier, Ohio State University, United States

S2S: SPEECH-TO-SPEECH TRANSLATION

S2S.1: MULTILINGUAL END-TO-END SPEECH TRANSLATION 570
Hirofumi Inaguma, Kyoto University, Japan; Kevin Duh, Johns Hopkins University, United States; Tatsuya Kawahara, Kyoto University, Japan; Shinji Watanabe, Johns Hopkins University, United States

S2S.2: NEURAL MACHINE TRANSLATION WITH ACOUSTIC EMBEDDING 578
Takatomo Kano, Nara Institute of Science and Technology, Japan; Sakriani Sakti, Satoshi Nakamura, Nara Institute of Science and Technology / RIKEN Center for Advanced Intelligence Project AIP, Japan

S2S.3: ONE-TO-MANY MULTILINGUAL END-TO-END SPEECH TRANSLATION 585
Mattia Antonino Di Gangi, Fondazione Bruno Kessler / University of Trento, Italy; Matteo Negri, Marco Turchi, Fondazione Bruno Kessler, Italy

S2S.4: SPEECH-TO-SPEECH TRANSLATION BETWEEN UNTRANSCRIBED UNKNOWN LANGUAGES 593
Andros Tjandra, Sakriani Sakti, Satoshi Nakamura, Nara Institute of Science and Technology, Japan

SDR: SPOKEN DOCUMENT RETRIEVAL

SDR.1: ENHANCED BERT-BASED RANKING MODELS FOR SPOKEN DOCUMENT RETRIEVAL 601
Hsiao-Yun Lin, Tien-Hong Lo, Berlin Chen, National Taiwan Normal University, Taiwan

SDR.2: VIRTUAL ADVERSARIAL TRAINING FOR DS-CNN BASED SMALL-FOOTPRINT KEYWORD SPOTTING 607
Xiong Wang, Northwestern Polytechnical University, China; Sining Sun, Northwestern Polytechnical University / Tencent, China; Lei Xie, Northwestern Polytechnical University, China

SDR.3: VERIFYING DEEP KEYWORD SPOTTING DETECTION WITH ACOUSTIC WORD EMBEDDINGS 613
Yougen Yuan, Northwestern Polytechnical University, China; Zhiqiang Lv, Shen Huang, Tencent Research, China; Lei Xie, Northwestern Polytechnical University, China

SDR.4: MULTILINGUAL BOTTLENECK FEATURES FOR QUERY BY EXAMPLE SPOKEN TERM DETECTION 621
Dhananjay Ram, Lesly Miculicich, Hervé Bourlard, Idiap Research Institute / École Polytechnique Fédérale de Lausanne, Switzerland

SDR.5: ADDITIONAL SHARED DECODER ON SIAMESE MULTI-VIEW ENCODERS FOR LEARNING ACOUSTIC WORD EMBEDDINGS 629
Myunghun Jung, Hyungjun Lim, Jahyun Goo, Youngmoon Jung, Hoirin Kim, Korea Advanced Institute of Science and Technology (KAIST), Korea (South)

SDR.6: EFFICIENT FREE KEYWORD DETECTION BASED ON CNN AND END-TO-END CONTINUOUS DP-MATCHING 637
Tomohiro Tanaka, Takahiro Shinozaki, Tokyo Institute of Technology, Japan

SDS: SPOKEN DIALOG SYSTEMS

SDS.1: ROBUST BELIEF STATE SPACE REPRESENTATION FOR STATISTICAL DIALOGUE MANAGERS USING DEEP AUTOENCODERS 1055
Fotios Lygerakis, Vassilios Diakoloulas, Michail Lagoudakis, Technical University of Crete, Greece; Kotti Margarita, Toshiba Research Cambridge, United Kingdom

SDS.2: IMPROVING SPEECH-BASED END-OF-TURN DETECTION VIA CROSS-MODAL□REPRESENTATION LEARNING WITH PUNCTUATED TEXT DATA 1062
Ryo Masumura, Mana Ihori, Tomohiro Tanaka, Atsushi Ando, Ryo Ishii, Takanobu Oba, Ryuichiro Higashinaka, NTT Corporation, Japan

SDS.3: DIALOGUE ENVIRONMENTS ARE DIFFERENT FROM GAMES: INVESTIGATING VARIANTS OF DEEP Q-NETWORKS FOR DIALOGUE POLICY 1070
Yu-An Wang, Yun-Nung Chen, National Taiwan University, Taiwan

SDS.4: EFFICIENT SEMI-SUPERVISED LEARNING FOR NATURAL LANGUAGE UNDERSTANDING BY OPTIMIZING DIVERSITY 1077
Eunah Cho, He Xie, Alexa AI, United States; John Lalor, University of Notre Dame, United States; Varun Kumar, William M. Campbell, Alexa AI, United States

SLR-1: SPEAKER/LANGUAGE RECOGNITION I

SLR-1.1: JOINT OPTIMIZATION OF CLASSIFICATION AND CLUSTERING FOR DEEP SPEAKER EMBEDDING 284
Zhiming Wang, Kaisheng Yao, Shuo Fang, Xiaolong Li, Ant Financial, Inc., China

SLR-1.2: EXPLORING EFFECTIVE DATA AUGMENTATION WITH TDNN-LSTM NEURAL NETWORK EMBEDDING FOR SPEAKER RECOGNITION 291
Chien-Lin Huang, PingAn AI Lab, United States

SLR-1.3: END-TO-END NEURAL SPEAKER DIARIZATION WITH SELF-ATTENTION 296
Yusuke Fujita, Naoyuki Kanda, Shota Horiguchi, Yawen Xue, Kenji Nagamatsu, Hitachi, Ltd., Japan; Shinji Watanabe, Johns Hopkins University, United States

SLR-1.4: A CROSS-CORPUS STUDY ON SPEECH EMOTION RECOGNITION 304
Rosanna Milner, Md Asif Jalal, University of Sheffield, United Kingdom; Raymond W. M. Ng, Emotech Labs, United Kingdom; Thomas Hain, University of Sheffield, United Kingdom

SLR-1.5: ADVERSARIAL ATTACKS ON SPOOFING COUNTERMEASURES OF AUTOMATIC SPEAKER VERIFICATION 312
Songxiang Liu, Chinese University of Hong Kong, China; Haibin Wu, Hung-yi Lee, National Taiwan University, Taiwan; Helen Meng, Chinese University of Hong Kong, China

SLR-1.6: SPOKEN LANGUAGE IDENTIFICATION USING BIDIRECTIONAL LSTM BASED LID SEQUENTIAL SENONES. 320
Muralikrishna H, Pulkit Sapra, Anuksha Jain, Dileep Aroor Dinesh, Indian Institute of Technology, Mandi, India

SLR-1.7: TIME-DOMAIN SPEAKER EXTRACTION NETWORK.....327
Chenglin Xu, Nanyang Technological University, Singapore; Wei Rao, National University of Singapore, Singapore; Eng Siong Chng, Nanyang Technological University, Singapore; Haizhou Li, National University of Singapore, Singapore

SLR-1.8: SHORT UTTERANCE COMPENSATION IN SPEAKER VERIFICATION VIA .....335 COSINE-BASED TEACHER-STUDENT LEARNING OF SPEAKER EMBEDDINGS
Jee-weon Jung, Hee-Soo Heo, Hye-jin Shim, Ha-Jin Yu, University of Seoul, Korea (South)

SLR-1.9: NOVEL ENHANCED TEAGER ENERGY BASED CEPSTRAL COEFFICIENTS .....342 FOR REPLAY SPOOF DETECTION
Rajul Acharya, Hemant A. Patil, Harsh Kotta, Dhirubhai Ambani Institute of Information and Communication Technology, India

SLR-1.10: SYLLABLE-DEPENDENT DISCRIMINATIVE LEARNING FOR SMALL .....350 FOOTPRINT TEXT-DEPENDENT SPEAKER VERIFICATION
Junyi Peng, Yuexian Zou, Peking University, China; Na Li, Deyi Tuo, Dan Su, Meng Yu, Chunlei Zhang, Dong Yu, Tencent, China

SLR-1.11: LATENT SPACE REPRESENTATION FOR MULTI-TARGET SPEAKER .....358 DETECTION AND IDENTIFICATION WITH A SPARSE DATASET USING TRIPLET NEURAL NETWORKS
Kin Wai Cheuk, Balamurali B T, Gemma Roig, Dorien Herremans, Singapore University of Technology and Design, Singapore

SLR-1.12: SELF-ADAPTIVE SOFT VOICE ACTIVITY DETECTION USING DEEP .....365 NEURAL NETWORKS FOR ROBUST SPEAKER VERIFICATION
Youngmoon Jung, Yeunju Choi, Hoirin Kim, Korea Advanced Institute of Science and Technology (KAIST), Korea (South)

SLR-1.13: SPHEREDIAR: AN EFFECTIVE SPEAKER DIARIZATION SYSTEM FOR .....373 MEETING DATA
Tuomas Kaseva, Aku Rouhe, Mikko Kurimo, Aalto University, Finland

SLR-1.14: BAYESIAN ADVERSARIAL LEARNING FOR SPEAKER RECOGNITION.....381
Jen-Tzung Chien, Chun-Lin Kuo, National Chiao Tung University, Taiwan

SLR-1.15: AN INVESTIGATION OF LSTM-CTC BASED JOINT ACOUSTIC MODEL .....389 FOR INDIAN LANGUAGE IDENTIFICATION
Tirusha Mandava, Ravi Kumar Vuddagiri, Hari Krishna Vydana, Anil Kumar Vuppala, IIIT Hyderabad, India

SLR-1.16: A MULTI PURPOSE AND LARGE SCALE SPEECH CORPUS IN PERSIAN AND .....397 ENGLISH FOR SPEAKER AND SPEECH RECOGNITION: THE DEEPMINE DATABASE
Hossein Zeinali, Lukáš Burget, Jan Cernocky, Brno University of Technology, Czech Republic

SLR-1.17: NATIVE LANGUAGE IDENTIFICATION FROM RAW WAVEFORMS USING .....403 DEEP CONVOLUTIONAL NEURAL NETWORKS WITH ATTENTIVE POOLING
Rutuja Ubale, Vikram Ramanarayanan, Yao Qian, Keelan Evanini, Chee Wee Leong, Chong Min Lee, Educational Testing Service, United States

SLR-1.18: SPEAKER VERIFICATION WITH APPLICATION-AWARE BEAMFORMING 411
Ladislav Mošner, Oldřich Plchot, Johan Rohdin, Lukáš Burget, Jan Černocký, Brno University of Technology, Czech Republic

SLR-2: SPEAKER/LANGUAGE RECOGNITION II

SLR-2.1: INCORPORATING PRIOR KNOWLEDGE INTO SPEAKER DIARIZATION AND .....697 LINKING FOR IDENTIFYING COMMON SPEAKER
Tsun-Yat Leung, Lahiru Samarakoon, Albert Y.S. Lam, Fano Labs, China

SLR-2.2: LOGISTIC SIMILARITY METRIC LEARNING VIA AFFINITY MATRIX FOR .....704 TEXT-INDEPENDENT SPEAKER VERIFICATION
Junyi Peng, Rongzhi Gu, Yuexian Zou, Peking University, China

SLR-2.3: LOW-RESOURCE DOMAIN ADAPTATION FOR SPEAKER RECOGNITION .....710 USING CYCLE-GANS
Phani Sankar Nidadavolu, Saurabh Kataria, Jesús Villalba, Najim Dehak, Johns Hopkins University, United States

SLR-2.4: CNN WITH PHONETIC ATTENTION FOR TEXT-INDEPENDENT SPEAKER .....718 VERIFICATION
Tianyan Zhou, Yong Zhao, Jinyu Li, Yifan Gong, Jian Wu, Microsoft Corporation, United States

SLR-2.5: PROBING THE INFORMATION ENCODED IN X-VECTORS.....726
Desh Raj, David Snyder, Daniel Povey, Sanjeev Khudanpur, The Johns Hopkins University, United States

SLR-2.6: IN-THE-WILD END-TO-END DETECTION OF SPEECH AFFECTING .....734 DISEASES
Joana Correia, Carnegie Mellon University / INESC-ID, Instituto Superior Tecnico / University of Lisbon, United States; Isabel Trancoso, INESC-ID, Instituto Superior Tecnico / University of Lisbon, Portugal; Bhiksha Raj, Carnegie Mellon University, United States

SLR-2.7: OPTIMIZING NEURAL NETWORK EMBEDDINGS USING A PAIR-WISE .....742 LOSS FOR TEXT-INDEPENDENT SPEAKER VERIFICATION
Hira Dhamyal, Tianyan Zhou, Bhiksha Raj, Rita Singh, Carnegie Mellon University, United States

SLR-2.8: TOWARDS CONTROLLING FALSE ALARM — MISS TRADE-OFF IN .....749 PERCEPTUAL SPEAKER COMPARISON VIA NON-NEUTRAL LISTENING TASK FRAMING
Rosa González Hautamäki, Tomi H. Kinnunen, University of Eastern Finland, Finland

SLR-2.9: DOVER: A METHOD FOR COMBINING DIARIZATION OUTPUTS.....757
Andreas Stolcke, Takuya Yoshioka, Microsoft, United States

SLU: SPOKEN LANGUAGE UNDERSTANDING

SLU.1: USING VERY DEEP CONVOLUTIONAL NEURAL NETWORKS TO .....764 AUTOMATICALLY DETECT PLAGIARIZED SPOKEN RESPONSES
Xinhao Wang, Keelan Evanini, Yao Qian, Klaus Zechner, Educational Testing Service, United States

SLU.2: SPOKEN MULTIPLE-CHOICE QUESTION ANSWERING USING MULTIMODAL CONVOLUTIONAL NEURAL NETWORKS 772
Shang-Bao Luo, National Taiwan University of Science and Technology, Taiwan; Hung-Shin Lee, Institute of Information Science, Academia Sinica, Taiwan; Kuan-Yu Chen, National Taiwan University of Science and Technology, Taiwan; Hsin-Min Wang, Institute of Information Science, Academia Sinica, Taiwan

SLU.3: TRANSFER LEARNING FOR CONTEXT-AWARE SPOKEN LANGUAGE UNDERSTANDING 779
Qian Chen, Zhu Zhuo, Wen Wang, Qiuyun Xu, Alibaba Group, China

SLU.4: EMOCEPTION: AN INCEPTION INSPIRED EFFICIENT SPEECH EMOTION RECOGNITION NETWORK 787
Chirag Singh, Abhay Kumar, Ajay Nagar, Suraj Tripathi, Promod Yenigalla, Samsung R&D Institute India - Bangalore, India

SLU.5: A COMPARATIVE STUDY ON END-TO-END SPEECH TO TEXT TRANSLATION 792
Parnia Bahar, Tobias Bieschke, Hermann Ney, RWTH Aachen University, Germany

SLU.6: JOINT LEARNING OF WORD AND LABEL EMBEDDINGS FOR SEQUENCE LABELLING IN SPOKEN LANGUAGE UNDERSTANDING 800
Jiewen Wu, Institute for Infocomm Research, Singapore; Luis Fernando D'Haro, Universidad Politecnica de Madrid, Spain; Nancy F. Chen, Pavitra Krishnaswamy, Institute for Infocomm Research, Singapore; Rafael Banchs, Intapp Inc., United States

SLU.7: MARKOV RECURRENT NEURAL NETWORK LANGUAGE MODEL....807
Jen-Tzung Chien, Che-Yu Kuo, National Chiao Tung University, Taiwan

SLU.8: TOPIC-AWARE POINTER-GENERATOR NETWORKS FOR SUMMARIZING SPOKEN CONVERSATIONS 814
*Zhengyuan Liu, Institute for Infocomm Research, A*STAR, Singapore; Angela Ng, Sheldon Lee, Changi General Hospital, Singapore; Ai Ti Aw, Nancy F. Chen, Institute for Infocomm Research, A*STAR, Singapore*

SLU.9: SLU FOR VOICE COMMAND IN SMART HOME: COMPARISON OF PIPELINE AND END-TO-END APPROACHES 822
Thierry Desot, François Portet, Michel Vacher, Laboratoire d'informatique de Grenoble, France

SLU.10: SCALABLE NEURAL DIALOGUE STATE TRACKING....830
Vevake Balaraman, Bernardo Magnini, Fondazione Bruno Kessler, Italy

SLU.11: HIERARCHICAL TRANSFORMERS FOR LONG DOCUMENT CLASSIFICATION 838
Raghavendra Pappagari, Johns Hopkins University, United States; Piotr Zelasko, Avaya Conversational Intelligence, United States; Jesus Villalba, Johns Hopkins Univeristy, United States; Yishay Carmiel, Avaya Conversational Intelligence, United States; Najim Dehak, Johns Hopkins University, United States

SLU.12: ADAPTING PRETRAINED TRANSFORMER TO LATTICES FOR SPOKEN LANGUAGE UNDERSTANDING 845
Chao-Wei Huang, Yun-Nung Chen, National Taiwan University, Taiwan

SLU.13: SPATIO-TEMPORAL CONTEXT MODELLING FOR SPEECH EMOTION CLASSIFICATION 853
Md Asif Jalal, Roger K. Moore, Thomas Hain, University of Sheffield, United Kingdom

SLU.14: PARAPHRASE GENERATION BASED ON VAE AND POINTER-GENERATOR NETWORKS 860
Lohith Ravuru, Hyungtak Choi, Siddarth K.M., Hojung Lee, Inchul Hwang, Samsung Electronics Co. Ltd., Korea (South)

SS-1: ASVSPOOF 2019: FUTURE HORIZONS IN SPOOFED AND FAKE AUDIO DETECTION, ANALYSING OPERATIONAL SETTINGS

SS-1.1: DEVELOPMENT OF VOICE SPOOFING DETECTION SYSTEMS FOR 2019 EDITION OF AUTOMATIC SPEAKER VERIFICATION AND COUNTERMEASURES CHALLENGE 1003
Joao Monteiro, Institut National de la Recherche Scientifique, Canada; Jahangir Alam, Centre de Recherche Informatique de Montreal, Canada

SS-1.2: SPOOF DETECTION USING TIME-DELAY SHALLOW NEURAL NETWORK AND FEATURE SWITCHING 1011
Mari Ganesh Kumar, Indian Institute of Technology, Madras, India; Suvidha Rupesh Kumar, SSN College of Engineering, India; Saranya M S, Indian Institute of Technology, Madras, India; Bharathi B, Shiva Subramaniya Nadar College of Engineering, India; Hema A Murthy, Indian Institute of Technology, Madras, India

SS-1.3: LONG RANGE ACOUSTIC AND DEEP FEATURES PERSPECTIVE ON ASVSPOOF 2019 1018
Rohan Kumar Das, Jichen Yang, Haizhou Li, National University of Singapore, Singapore

SS-2: MGB-5: THE FIFTH EDITION OF THE MULTI-GENRE BROADCAST CHALLENGE

SS-2.1: THE MGB-5 CHALLENGE: RECOGNITION AND DIALECT IDENTIFICATION OF DIALECTAL ARABIC SPEECH 1026
Ahmed Ali, Hamad Bin Khalifa University, Qatar; Suwon Shon, Massachusetts Institute of Technology, United States; Younes Samih, Hamdy Mubarak, Ahmed Abdelali, Hamad Bin Khalifa University, Qatar; James Glass, Massachusetts Institute of Technology, United States; Steve Renals, University of Edinburgh, United Kingdom; Khalid Choukri, European Language Resources Association, France

SS-2.2: RDI-CU SYSTEM FOR THE 2019 ARABIC MULTI-GENRE BROADCAST CHALLENGE .....N/A
Hany Ahmed, Hazem Mamdouh, Salah Ashraf, Ali Ramadan, Mohsen Rashwan, Research and Development International Company, Egypt

TTS: TEXT-TO-SPEECH SYSTEM

TTS.1: ON THE STUDY OF GENERATIVE ADVERSARIAL NETWORKS FOR CROSS-LINGUAL VOICE CONVERSION 144
*Berrak Sisman, Mingyang Zhang, National University of Singapore, Singapore; Minghui Dong, Institute for Infocomm Research, A*STAR, Singapore; Haizhou Li, National University of Singapore, Singapore*

TTS.2: WAVENET FACTORIZATION WITH SINGULAR VALUE DECOMPOSITION FOR VOICE CONVERSION 152
Hongqiang Du, Northwestern Polytechnical University, China; Xiaohai Tian, National University of Singapore, Singapore; Lei Xie, Northwestern Polytechnical University, China; Haizhou Li, National University of Singapore, Singapore

TTS.3: A MODULARIZED NEURAL NETWORK WITH LANGUAGE-SPECIFIC OUTPUT LAYERS FOR CROSS-LINGUAL VOICE CONVERSION 160
Yi Zhou, Xiaohai Tian, Emre Yilmaz, Rohan Kumar Das, Haizhou Li, National University of Singapore, Singapore

TTS.4: KNOWLEDGE DISTILLATION FROM BERT IN PRE-TRAINING AND FINE-TUNING FOR POLYPHONE DISAMBIGUATION 168
Hao Sun, Peking University, China; Xu Tan, Microsoft Research, China; Jun-Wei Gan, Sheng Zhao, Dongxu Han, Microsoft STC Asia, China; Hongzhi Liu, Peking University, China; Tao Qin, Tie-Yan Liu, Microsoft Research, China

TTS.5: INVESTIGATION OF SHALLOW WAVENET VOCODER WITH LAPLACIAN DISTRIBUTION OUTPUT 176
Patrick Lumban Tobing, Tomoki Hayashi, Tomoki Toda, Nagoya University, Japan

TTS.6: LEARNING HIERARCHICAL REPRESENTATIONS FOR EXPRESSIVE SPEAKING STYLE IN END-TO-END SPEECH SYNTHESIS 184
Xiaochun An, Northwestern Polytechnical University, China; Yuxuan Wang, ByteDance AI Lab, China; Shan Yang, Northwestern Polytechnical University, China; Zejun Ma, ByteDance AI Lab, China; Lei Xie, Northwestern Polytechnical University, China

TTS.7: CONTROLLING EMOTION STRENGTH WITH RELATIVE ATTRIBUTE FOR END-TO-END SPEECH SYNTHESIS 192
Xiaolian Zhu, Shan Yang, Geng Yang, Lei Xie, Northwestern Polytechnical University, China

TTS.8: BOOTSTRAPPING NON-PARALLEL VOICE CONVERSION FROM SPEAKER-ADAPTIVE TEXT-TO-SPEECH 200
Hieu-Thi Luong, Junichi Yamagishi, National Institute of Informatics, Japan

TTS.9: IMPROVING MANDARIN END-TO-END SPEECH SYNTHESIS BY SELF-ATTENTION AND LEARNABLE GAUSSIAN BIAS 208
Fengyu Yang, Shan Yang, Northwestern Polytechnical University, China; Pengcheng Zhu, Pengju Yan, Tongdun, China; Lei Xie, Northwestern Polytechnical University, China

TTS.10: TACOTRON-BASED ACOUSTIC MODEL USING PHONEME ALIGNMENT FOR PRACTICAL NEURAL TEXT-TO-SPEECH SYSTEMS 214
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