

34th Conference on Computational Linguistics and Speech Processing (ROCLING 2022)

Taipei, Taiwan
21-22 November 2022

Editors:

**Yung-Chun Chang
Yi-Chin Huang
Jheng-Long Wu
Ming-Hsiang Su
Hen-Hsen Huang**

**Yi-Fen Liu
Lung-Hao Lee
Chin-Hung Chou
Yuan-Fu Liao**

ISBN: 978-1-7138-6764-7

Printed from e-media with permission by:

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



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

Copyright© (2022) by the Association for Computational Linguistics and Chinese Language Processing (ACLCLP), Authors of Papers.
All rights reserved.

Printed with permission by Curran Associates, Inc. (2023)

For permission requests, please contact the Association for Computational Linguistics at the address below.

Association for Computational Linguistics
209 N. Eighth Street
Stroudsburg, Pennsylvania 18360

Phone: 1-570-476-8006
Fax: 1-570-476-0860

acl@aclweb.org

Additional copies of this publication are available from:

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

Table of Contents

Language Model Based Chinese Handwriting Address Recognition	1
Chinese Movie Dialogue Question Answering Dataset	7
Unsupervised Text Summarization of Long Documents using Dependency-based Noun Phrases and Contextual Order Arrangement	15
Enhancing Chinese Multi-Label Text Classification Performance with Response-based Knowledge Distillation	25
Combining Word Vector Technique and Clustering Algorithm for Credit Card Merchant Detection	32
Taiwanese-Accented Mandarin and English Multi-Speaker Talking-Face Synthesis System	40
Is Character Trigram Overlapping Ratio Still the Best Similarity Measure for Aligning Sentences in a Paraphrased Corpus?	49
RoBERTa-based Traditional Chinese Medicine Named Entity Recognition Model	61
A Study on Using Different Audio Lengths in Transfer Learning for Improving Chainsaw Sound Recognition	67
Using Grammatical and Semantic Correction Model to Improve Chinese-to-Taiwanese Machine Translation Fluency	75
Investigation of feature processing modules and attention mechanisms in speaker verification system	84
A Preliminary Study of the Application of Discrete Wavelet Transform Features in Conv-TasNet Speech Enhancement Model	92

Exploiting the compressed spectral loss for the learning of the DEMUCS speech enhancement network	100
Using Machine Learning and Pattern-Based Methods for Identifying Elements in Chinese Judgment Documents of Civil Cases	107
Development of Mandarin-English code-switching speech synthesis system	116
Predicting Judgments and Grants for Civil Cases of Alimony for the Elderly	121
Lightweight Sound Event Detection Model with RepVGG Architecture	129
Analyzing discourse functions with acoustic features and phone embeddings: non-lexical items in Taiwan Mandarin	136
A Dimensional Valence-Arousal-Irony Dataset for Chinese Sentence and Context	147
Intelligent Future Recreation Harbor Application Service: Taking Kaohsiung Asia New Bay as an Example to Construct a Composite Recreational Knowledge Graph	155
HanTrans: An Empirical Study on Cross-Era Transferability of Chinese Pre-trained Language Model	164
A Preliminary Study on Automated Speaking Assessment of English as a Second Language (ESL) Students	174
Clustering Issues in Civil Judgments for Recommending Similar Cases	184
Multifaceted Assessments of Traditional Chinese Word Segmentation Tool on Large Corpora	193
Mandarin-English Code-Switching Speech Recognition System for Specific Domain	200

Legal Case Winning Party Prediction With Domain Specific Auxiliary Models	205
Early Speech Production in Infants and Toddlers Later Diagnosed with Cerebral Palsy: A Retrospective Study	214
Automatic Generation of Abstracts for Research Papers	221
Speech Timing in Typically Developing Mandarin-Speaking Children From Ages 3 To 4	230
Right-Dominant Tones in Zhangzhou: On and Through Phonetic Surface	236
Web-API-Based Chatbot Generation with Analysis and Expansion for Training Sentences	246
The Design and Development of a System for Chinese Character Difficulty and Features	256
Image Caption Generation for Low-Resource Assamese Language	263
Building an Enhanced Autoregressive Document Retriever Leveraging Supervised Contrastive Learning	273
A Quantitative Analysis of Comparison of Emoji Sentiment: Taiwan Mandarin Users and English Users	283
Applying Information Extraction to Storybook Question and Answer Generation	289
Improving Response Diversity through Commonsense-Aware Empathetic Response Generation	299
A Preliminary Study on Mandarin-Hakka neural machine translation using small-sized data	307

NCU1415 at ROCLING 2022 Shared Task: A light-weight transformer-based approach for Biomedical Name Entity Recognition	316
CrowNER at Rocling 2022 Shared Task: NER using MacBERT and Adversarial Training	321
SCU-MESCLab at ROCLING-2022 Shared Task: Named Entity Recognition Using BERT Classifier	329
YNU-HPCC at ROCLING 2022 Shared Task: A Transformer-based Model with Focal Loss and Regularization Dropout for Chinese Healthcare Named Entity Recognition	335
NERVE at ROCLING 2022 Shared Task: A Comparison of Three Named Entity Recognition Frameworks Based on Language Model and Lexicon Approach	343
SCU-NLP at ROCLING 2022 Shared Task: Experiment and Error Analysis of Biomedical Entity Detection Model	350
MIGBaseline at ROCLING 2022 Shared Task: Report on Named Entity Recognition Using Chinese Healthcare Datasets	356
Overview of the ROCLING 2022 Shared Task for Chinese Healthcare Named Entity Recognition	363