

Tan Lee

List of Publications by Year in descending order

Source: <https://exaly.com/author-pdf/5788532/publications.pdf>

Version: 2024-02-01

160
papers

1,436
citations

516561

16
h-index

580701

25
g-index

162
all docs

162
docs citations

162
times ranked

833
citing authors

#	ARTICLE	IF	CITATIONS
1	Enhancing Segment-Based Speech Emotion Recognition by Iterative Self-Learning. IEEE/ACM Transactions on Audio Speech and Language Processing, 2022, 30, 123-134.	4.0	13
2	Bayesian Learning for Deep Neural Network Adaptation. IEEE/ACM Transactions on Audio Speech and Language Processing, 2021, 29, 2096-2110.	4.0	9
3	Estimating Mutual Information in Prosody Representation for Emotional Prosody Transfer in Speech Synthesis. , 2021, , .		4
4	Low-Resource NMT: A Case Study on the Written and Spoken Languages in Hong Kong. , 2021, , .		1
5	Automatic Assessment of Speech Impairment in Cantonese-Speaking People with Aphasia. IEEE Journal on Selected Topics in Signal Processing, 2020, 14, 331-345.	7.3	13
6	Mixture Factorized Auto-Encoder for Unsupervised Hierarchical Deep Factorization of Speech Signal. , 2020, , .		2
7	Time-Frequency Feature Decomposition Based on Sound Duration for Acoustic Scene Classification. , 2020, , .		10
8	Resting-State EEG-Based Biometrics with Signals Features Extracted by Multivariate Empirical Mode Decomposition. , 2020, , .		0
9	An End-to-End Approach to Automatic Speech Assessment for Cantonese-speaking People with Aphasia. Journal of Signal Processing Systems, 2020, 92, 819-830.	1.4	14
10	Exploiting Cross-Lingual Speaker and Phonetic Diversity for Unsupervised Subword Modeling. IEEE/ACM Transactions on Audio Speech and Language Processing, 2019, 27, 2000-2011.	4.0	8
11	Combining Phone Posteriorgrams from Strong and Weak Recognizers for Automatic Speech Assessment of People with Aphasia. , 2019, , .		3
12	Adversarial Multi-task Deep Features and Unsupervised Back-end Adaptation for Language Recognition. , 2019, , .		4
13	Enhancing Sound Texture in CNN-based Acoustic Scene Classification. , 2019, , .		21
14	BLHUC: Bayesian Learning of Hidden Unit Contributions for Deep Neural Network Speaker Adaptation. , 2019, , .		16
15	Acoustical Assessment of Voice Disorder With Continuous Speech Using ASR Posterior Features. IEEE/ACM Transactions on Audio Speech and Language Processing, 2019, 27, 1047-1059.	4.0	19
16	Revisiting Hidden Markov Models for Speech Emotion Recognition. , 2019, , .		41
17	Guest Editorial: Advances in Deep Learning for Speech Processing. Journal of Signal Processing Systems, 2018, 90, 959-961.	1.4	4
18	Disordered Speech Assessment Using Kullback-Leibler Divergence Features with Multi-Task Acoustic Modeling. , 2018, , .		4

#	ARTICLE	IF	CITATIONS
19	Investigation of Stacked Deep Neural Networks and Mixture Density Networks for Acoustic-to-Articulatory Inversion. , 2018, , .		3
20	An End-to-End Approach to Automatic Speech Assessment for People with Aphasia. , 2018, , .		4
21	Prediction of Voice Disorder Severity: Contributions from Sustained Vowels and Continuous Speech. , 2018, , .		2
22	A Study on Acoustic Modeling for Child Speech Based on Multi-Task Learning. , 2018, , .		5
23	An Automated Assessment Tool for Child Speech Disorders. , 2018, , .		2
24	Unsupervised Pattern Discovery from Thematic Speech Archives Based on Multilingual Bottleneck Features. , 2018, , .		1
25	Automatic Speech Assessment for Aphasic Patients Based on Syllable-Level Embedding and Supra-Segmental Duration Features. , 2018, , .		10
26	Reducing Model Complexity for DNN Based Large-Scale Audio Classification. , 2018, , .		15
27	Improving Mobile Phone Speech Recognition by Personalized Amplification: Application in People with Normal Hearing and Mild-to-Moderate Hearing Loss. Ear and Hearing, 2017, 38, e85-e92.	1.0	8
28	Polyphonic piano note transcription with non-negative matrix factorization of differential spectrogram. , 2017, , .		3
29	Shefce: A Cantonese-English bilingual speech corpus for pronunciation assessment. , 2017, , .		3
30	Audio-visual expressions of attitude: How many different attitudes can perceivers decode?. Speech Communication, 2017, 95, 114-126.	1.6	10
31	Exploiting language-mismatched phoneme recognizers for unsupervised acoustic modeling. , 2016, , .		4
32	Towards automatic assessment of aphasia speech using automatic speech recognition techniques. , 2016, , .		5
33	Automatic speech recognition for acoustical analysis and assessment of cantonese pathological voice and speech. , 2016, , .		24
34	Surface Electromyographic Activity of Extrinsic Laryngeal Muscles in Cantonese Tone Production. Journal of Signal Processing Systems, 2016, 82, 287-294.	1.4	7
35	Analysis of intonation patterns in Cantonese aphasia speech. , 2015, 2015, 86-89.		1
36	Multi-pitch estimation based on sparse representation with pre-screened dictionary. , 2015, , .		1

#	ARTICLE	IF	CITATIONS
37	A method of speech periodicity enhancement using transform-domain signal decomposition. <i>Speech Communication</i> , 2015, 67, 102-112.	1.6	11
38	Acoustic Segment Modeling with Spectral Clustering Methods. <i>IEEE/ACM Transactions on Audio Speech and Language Processing</i> , 2015, 23, 264-277.	4.0	47
39	Objective measures for quality assessment of noise-suppressed speech. <i>Speech Communication</i> , 2015, 71, 62-73.	1.6	7
40	Supervised Single-Microphone Multi-Talker Speech Separation with Conditional Random Fields. <i>IEEE/ACM Transactions on Audio Speech and Language Processing</i> , 2015, 23, 2334-2342.	4.0	2
41	The effect of enhancing temporal periodicity cues on Cantonese tone recognition by cochlear implantees. <i>International Journal of Audiology</i> , 2014, 53, 546-557.	0.9	5
42	Surface electromyographic activity of non-laryngeal neck muscles in Cantonese tone production. , 2014, , .		3
43	Multipitch tracking based on linear programming relaxation and sparsity-based pitch candidate estimation. , 2014, , .		1
44	Improving the sound quality of an electronic voice box. , 2014, , .		1
45	Correcting Chord Classification Errors Based on Tonal Organization Information of Classical Music. , 2014, , .		0
46	Spoken Language Recognition With Prosodic Features. <i>IEEE Transactions on Audio Speech and Language Processing</i> , 2013, 21, 1841-1853.	3.8	19
47	Chord classification of multi-instrumental music using exemplar-based sparse representation. , 2013, , .		1
48	Improving the sound quality of an electronic voice box. , 2013, , .		0
49	Using dynamic conditional random field on single-microphone speech separation. , 2013, , .		3
50	Shifted-Delta MLP Features for Spoken Language Recognition. <i>IEEE Signal Processing Letters</i> , 2013, 20, 15-18.	2.1	37
51	Analysis of Auto-aligned and Auto-segmented Oral Discourse by Speakers with Aphasia: A Preliminary Study on the Acoustic Parameter of Duration. <i>Procedia, Social and Behavioral Sciences</i> , 2013, 94, 71-72.	0.5	8
52	Pitch Estimation in Noisy Speech Using Accumulated Peak Spectrum and Sparse Estimation Technique. <i>IEEE Transactions on Audio Speech and Language Processing</i> , 2013, 21, 99-109.	3.8	36
53	Structured mean field method for single-microphone speech separation with factorial Hidden Markov Model. , 2013, , .		0
54	Using parallel tokenizers with DTW matrix combination for low-resource spoken term detection. , 2013, , .		32

#	ARTICLE	IF	CITATIONS
55	A speech enhancement method for cochlear implant listeners. , 2013, 2013, 2036-9.		0
56	Evaluation of pitch estimation algorithms on separated speech. , 2013, , .		2
57	Sparsity-based confidence measure for pitch estimation in noisy speech. , 2012, , .		6
58	An acoustic segment modeling approach to query-by-example spoken term detection. , 2012, , .		44
59	Transform-domain Wiener filter for speech periodicity enhancement. , 2012, , .		6
60	Two objective measures for speech distortion and noise reduction evaluation of enhanced speech signals. , 2012, , .		1
61	Clinical evaluation of a computerized self-administered hearing test. International Journal of Audiology, 2012, 51, 606-610.	0.9	33
62	Integrating multiple observations for model-based single-microphone speech separation with conditional random fields. , 2012, , .		4
63	Exploration of Phase and Vocal Excitation Modulation Features for Speaker Recognition. Lecture Notes in Computer Science, 2012, , 251-259.	1.0	0
64	Unsupervised spoken term detection with acoustic segment model. , 2011, , .		20
65	Dealing with imperfections in human speech communication with advanced speech processing techniques. , 2011, , .		1
66	Model-based non-negative matrix factorization for single-channel speech separation. , 2011, , .		1
67	Robust Speaker Recognition Using Denoised Vocal Source and Vocal Tract Features. IEEE Transactions on Audio Speech and Language Processing, 2011, 19, 196-205.	3.8	41
68	Transform-domain speech periodicity enhancement with adaptive coefficient weighting. , 2011, , .		2
69	Score fusion and calibration in multiple language detectors with large performance variation. , 2011, , .		0
70	Prosodic attribute model for spoken language identification. , 2010, , .		13
71	Improved Cantonese Tone Recognition with Approximated F0 Contour: Implications for Cochlear Implants. , 2010, , .		0
72	Semantics-based language modeling for Cantonese-English code-mixing speech recognition. , 2010, , .		11

#	ARTICLE	IF	CITATIONS
73	Robust speaker verification using phase information of speech. , 2010, , .		0
74	Spectral trajectory estimation using nonnegative matrix factorization for model-based monaural speech separation. , 2010, , .		0
75	SURE-MSE speech enhancement for robust speech recognition. , 2010, , .		7
76	Perception and analysis of linearly approximated F0 contours in Cantonese speech. , 2010, , .		0
77	Robust Speaker Recognition Using Denoised Vocal. IEEE Transactions on Audio Speech and Language Processing, 2010, , .	3.8	0
78	Analysis and Selection of Prosodic Features for Language Identification. , 2009, , .		13
79	Cantonese tone recognition with enhanced temporal periodicity cues. Journal of the Acoustical Society of America, 2009, 126, 327-337.	0.5	23
80	Development of the computerized Cantonese Disyllabic Lexical Tone Identification Test in Noise (CANDILET-N). Cochlear Implants International, 2009, 10, 130-137.	0.5	1
81	Cantonese lexical tone recognition from frequency-specific temporal envelope and periodicity components in the same versus different noise band carriers. Cochlear Implants International, 2009, , n/a-n/a.	0.5	0
82	Analysis and Signal Processing of Oesophageal and Pathological Voices. Eurasip Journal on Advances in Signal Processing, 2009, 2009, .	1.0	6
83	Tone-enhanced generalized character posterior probability (GCPP) for Cantonese LVCSR. Computer Speech and Language, 2008, 22, 360-373.	2.9	6
84	A Perceptual Study of Approximated Cantonese Tone Contours. , 2008, , .		2
85	Prosodic Variation in Cantonese-English Code-Mixed Speech. , 2008, , .		1
86	Deriving MFCC Parameters from the Dynamic Spectrum for Robust Speech Recognition. , 2008, , .		3
87	Entropy-Based Analysis of the Prosodic Features of Chinese Dialects. , 2008, , .		3
88	Pitch Tracking for Model-Based Speech Separation. , 2008, , .		2
89	Mandarin Tone Perception with Temporal Envelope and Periodicity Cues from Different Frequency Regions. , 2008, , .		4
90	Integration of Complementary Acoustic Features for Speaker Recognition. IEEE Signal Processing Letters, 2007, 14, 181-184.	2.1	65

#	ARTICLE	IF	CITATIONS
91	Robust Speaker Recognition Using Both Vocal Source and Vocal Tract Features Estimated from Noisy Input Utterances. , 2007, , .		3
92	Band-specific temporal periodicity enhancement for Cantonese tone perception with noise-excited vocoder. Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2007, 2007, 694-7.	0.5	4
93	Tone recognition in continuous Cantonese speech using supratone models. Journal of the Acoustical Society of America, 2007, 121, 2936-2945.	0.5	15
94	Frequency-Specific Temporal Envelope and Periodicity Components for Lexical Tone Identification in Cantonese. Ear and Hearing, 2007, 28, 107S-113S.	1.0	8
95	A power-based adaptive method for eigenanalysis without square-root operations. , 2007, 17, 209-224.		3
96	Static and Dynamic Spectral Features: Their Noise Robustness and Optimal Weights for ASR. IEEE Transactions on Audio Speech and Language Processing, 2007, 15, 1087-1097.	3.8	5
97	Discrimination Power of Vocal Source and Vocal Tract Related Features for Speaker Segmentation. IEEE Transactions on Audio Speech and Language Processing, 2007, 15, 1884-1892.	3.8	31
98	Speech recognition on DSP: issues on computational efficiency and performance analysis. Microprocessors and Microsystems, 2006, 30, 155-164.	1.8	18
99	CANTONESE SPEECH RECOGNITION AND SYNTHESIS. , 2006, , 365-386.		10
100	TONE MODELING FOR SPEECH RECOGNITION. , 2006, , 179-200.		6
101	Analysis and modeling of F0 contours for cantonese text-to-speech. ACM Transactions on Asian Language Information Processing, 2004, 3, 169-180.	0.8	33
102	Robust adaptive quasi-Newton algorithms for eigensubspace estimation. IET Computer Vision, 2003, 150, 321.	1.3	17
103	A real-time Chinese speech recognition IC with double mixtures. , 2003, , .		2
104	Using tone information in Cantonese continuous speech recognition. ACM Transactions on Asian Language Information Processing, 2002, 1, 83-102.	0.8	38
105	Quasi-Newton algorithm for adaptive minor component extraction. Electronics Letters, 2002, 38, 1142.	0.5	7
106	A new approach to generating Pitch Cycle Waveform (PCW) for Waveform Interpolation codec. Microprocessors and Microsystems, 2002, 25, 421-426.	1.8	1
107	Spoken language resources for Cantonese speech processing. Speech Communication, 2002, 36, 327-342.	1.6	76
108	Noisy speech recognition using de-noised multiresolution analysis acoustic features. Journal of the Acoustical Society of America, 2001, 110, 2567-2574.	0.5	8

#	ARTICLE	IF	CITATIONS
109	A Low Missing Rate Audio Search Technique for Cantonese Radio Broadcast Recording. Lecture Notes in Computer Science, 2001, , 546-549.	1.0	0
110	Two-dimensional multi-resolution analysis of speech signals and its application to speech recognition. , 1999, , .		5
111	Cantonese syllable recognition using neural networks. IEEE Transactions on Speech and Audio Processing, 1999, 7, 466-472.	2.0	15
112	Isolated word recognition using modular recurrent neural networks. Pattern Recognition, 1998, 31, 751-760.	5.1	18
113	Tone recognition of isolated Cantonese syllables. IEEE Transactions on Speech and Audio Processing, 1995, 3, 204-209.	2.0	47
114	An NN based tone classifier for Cantonese. , 0, , .		2
115	Recurrent neural networks for speech modeling and speech recognition. , 0, , .		8
116	Development of a large vocabulary speech database for Cantonese. , 0, , .		3
117	A neural network based speech recognition system for isolated Cantonese syllables. , 0, , .		4
118	An HMM-based speech recognition IC. , 0, , .		13
119	Cantonese verbal information verification system using GMM-based anti-model. , 0, , .		1
120	On noise robustness of dynamic and static features for continuous Cantonese digit recognition. , 0, , .		2
121	Detection of language boundary in code-switching utterances by bi-phone probabilities. , 0, , .		22
122	Static and Dynamic Spectral Features: Their Noise Robustness and Optimal Weights for ASR. , 0, , .		5
123	Speech recognition on DSP: issues on computational efficiency and performance analysis. , 0, , .		7
124	Feature Extraction From Talking Mouths for Video-Based Bi-Modal Speaker Verification. , 0, , .		0
125	Use of Vocal Source Features in Speaker Segmentation. , 0, , .		4
126	Predicting Severity of Voice Disorder from DNN-HMM Acoustic Posteriors. , 0, , .		9

#	ARTICLE	IF	CITATIONS
127	Automatic Speech Assessment for People with Aphasia Using TDNN-BLSTM with Multi-Task Learning. , 0, , .		17
128	Improving Unsupervised Subword Modeling via Disentangled Speech Representation Learning and Transformation. , 0, , .		6
129	Fast DNN Acoustic Model Speaker Adaptation by Learning Hidden Unit Contribution Features. , 0, , .		4
130	Tone information as a confidence measure for improving Cantonese LVCSR. , 0, , .		1
131	Development of a Cantonese-English code-mixing speech corpus. , 0, , .		20
132	Perceptual equivalence of approximated Cantonese tone contours. , 0, , .		5
133	Modeling tones in hakka on the basis of the command-response model. , 0, , .		0
134	Language modeling for speech recognition of spoken Cantonese. , 0, , .		3
135	Prosody for Mandarin speech recognition: a comparative study of read and spontaneous speech. , 0, , .		2
136	Exploration of vocal excitation modulation features for speaker recognition. , 0, , .		4
137	Effects of language mixing for automatic recognition of Cantonese-English code-mixing utterances. , 0, , .		1
138	Model-based speech separation: identifying transcription using orthogonality. , 0, , .		1
139	Towards long-range prosodic attribute modeling for language recognition. , 0, , .		2
140	Pitch estimation in noisy speech based on temporal accumulation of spectrum peaks. , 0, , .		9
141	Perception-based automatic approximation of F0 contours in Cantonese speech. , 0, , .		0
142	Exploitation of phase information for speaker recognition. , 0, , .		3
143	Robust pitch estimation using l1-regularized maximum likelihood estimation. , 0, , .		6
144	Unsupervised mining of acoustic subword units with segment-level Gaussian posteriorgrams. , 0, , .		14

#	ARTICLE	IF	CITATIONS
145	A graph-based Gaussian component clustering approach to unsupervised acoustic modeling. , 0, , .		7
146	Large-margin conditional random fields for single-microphone speech separation. , 0, , .		0
147	Modeling temporal dependency for robust estimation of LP model parameters in speech enhancement. , 0, , .		0
148	Hybrid Accelerated Optimization for Speech Recognition. , 0, , .		1
149	RNN-LDA Clustering for Feature Based DNN Adaptation. , 0, , .		3
150	On the Linguistic Relevance of Speech Units Learned by Unsupervised Acoustic Modeling. , 0, , .		2
151	Acoustic Assessment of Disordered Voice with Continuous Speech Based on Utterance-Level ASR Posterior Features. , 0, , .		10
152	Information structure and prosodic prominence: how does sentence final particle affect Cantonese intonation?. , 0, , .		0
153	Free Labeling of Audio-visual Attitudinal Expressions in Cantonese. , 0, , .		0
154	Exploiting Speaker and Phonetic Diversity of Mismatched Language Resources for Unsupervised Subword Modeling. , 0, , .		5
155	Cross-cultural (A)symmetries in Audio-visual Attitude Perception. , 0, , .		0
156	Improving Cross-Lingual Knowledge Transferability Using Multilingual TDNN-BLSTM with Language-Dependent Pre-Final Layer. , 0, , .		4
157	Combining Adversarial Training and Disentangled Speech Representation for Robust Zero-Resource Subword Modeling. , 0, , .		8
158	Deep Learning of Segment-Level Feature Representation with Multiple Instance Learning for Utterance-Level Speech Emotion Recognition. , 0, , .		14
159	Automatic Assessment of Language Impairment Based on Raw ASR Output. , 0, , .		2
160	Emotion Profile Refinery for Speech Emotion Classification. , 0, , .		2