Yonghong Yan

List of Publications by Year
in descending order

148 papers	710 citations	1040056 9 h-index	940533 16 g-index
151 all docs	151 docs citations	151 times ranked	427 citing authors

#	Article	IF	Citations
1	Source localization using deep neural networks in a shallow water environment. Journal of the Acoustical Society of America, 2018, 143, 2922-2932.	1.1	86
2	An Exploration of Dropout with LSTMs. , 0, , .		57
3	Transformer-Based Online CTC/Attention End-To-End Speech Recognition Architecture. , 2020, , .		53
4	Online Hybrid CTC/Attention End-to-End Automatic Speech Recognition Architecture. IEEE/ACM Transactions on Audio Speech and Language Processing, 2020, 28, 1452-1465.	5.8	36
5	Markovian discriminative modeling for cross-domain dialog state tracking. , 2014, , .		30
6	An Acoustic Traffic Monitoring System: Design and Implementation. , 2015, , .		18
7	A fast two-microphone noise reduction algorithm based on power level ratio for mobile phone. , 2012, , .		15
8	Window-Dominant Signal Subspace Methods for Multiple Short-Term Speech Source Localization. IEEE/ACM Transactions on Audio Speech and Language Processing, 2017, 25, 731-744.	5.8	15
9	Robust and Fast Localization of Single Speech Source Using a Planar Array. IEEE Signal Processing Letters, 2013, 20, 909-912.	3.6	14
10	Oracle performance investigation of the ideal masks. , 2016, , .		14
11	Mandarin-English bilingual Speech Recognition for real world music retrieval. Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, 2008, , .	1.8	10
12	Subsample time delay estimation via improved GCC PHAT algorithm. , 2008, , .		9
13	Automatic Piano Music Transcription Using Audioâ€Visual Features. Chinese Journal of Electronics, 2015, 24, 596-603.	1.5	9
14	Multi-Task Learning in Deep Neural Networks for Mandarin-English Code-Mixing Speech Recognition. IEICE Transactions on Information and Systems, 2016, E99.D, 2554-2557.	0.7	9
15	Target Speaker Localization Based on the Complex Watson Mixture Model and Time-Frequency Selection Neural Network. Applied Sciences (Switzerland), 2018, 8, 2326.	2.5	9
16	Multiple Source Localization in a Shallow Water Waveguide Exploiting Subarray Beamforming and Deep Neural Networks. Sensors, 2019, 19, 4768.	3.8	9
17	FSCNet: Feature-Specific Convolution Neural Network for Real-Time Speech Enhancement. IEEE Signal Processing Letters, 2021, 28, 1958-1962.	3.6	9
18	A Secondary Path-Decoupled Active Noise Control Algorithm Based on Deep Learning. IEEE Signal Processing Letters, 2022, 29, 234-238.	3.6	9

#	Article	IF	Citations
19	ETEH: Unified Attention-Based End-to-End ASR and KWS Architecture. IEEE/ACM Transactions on Audio Speech and Language Processing, 2022, 30, 1360-1373.	5.8	9
20	Evaluation of objective intelligibility prediction measures for noise-reduced signals in mandarin. , 2012, , .		8
21	Agglutinative Language Speech Recognition Using Automatic Allophone Deriving. Chinese Journal of Electronics, 2016, 25, 328-333.	1.5	8
22	Using Highway Connections to Enable Deep Smallâ€footprint LSTMâ€RNNs for Speech Recognition. Chinese Journal of Electronics, 2019, 28, 107-112.	1.5	8
23	The Design of Backend Classifiers in PPRLM System for Language Identification. , 2007, , .		7
24	A Computer-Assist Algorithm to Detect Repetitive Stuttering Automatically. , 2013, , .		7
25	Customer voice sensor: A comprehensive opinion mining system for call center conversation. , 2016, , .		7
26	A Model Compression Method With Matrix Product Operators for Speech Enhancement. IEEE/ACM Transactions on Audio Speech and Language Processing, 2020, 28, 2837-2847.	5.8	7
27	Pre-Training Transformer Decoder for End-to-End ASR Model with Unpaired Text Data. , 2021, , .		7
28	A two-microphone based voice activity detection for distant-talking speech in wide range of direction of arrival. , 2012, , .		6
29	Automatic Vocal Segments Detection in Popular Music. , 2013, , .		6
30	Information Fusion in Automatic User Satisfaction Analysis in Call Center. , 2016, , .		6
31	Robust multiple speech source localization using time delay histogram. , 2016, , .		6
32	Polyphonic Piano Transcription with a Note-Based Music Language Model. Applied Sciences (Switzerland), 2018, 8, 470.	2.5	6
33	TEnet: target speaker extraction network with accumulated speaker embedding for automatic speech recognition. Electronics Letters, 2019, 55, 816-819.	1.0	6
34	Keyword Search Using Attention-Based End-to-End ASR and Frame-Synchronous Phoneme Alignments. IEEE/ACM Transactions on Audio Speech and Language Processing, 2021, 29, 3202-3215.	5.8	6
35	Alleviating ASR Long-Tailed Problem by Decoupling the Learning of Representation and Classification. IEEE/ACM Transactions on Audio Speech and Language Processing, 2022, 30, 340-354.	5.8	6
36	An Effective Automated Essay Scoring System Using Support Vector Regression. , 2012, , .		5

#	Article	IF	Citations
37	Parallel implementation of neural networks training on graphic processing unit. , 2012, , .		5
38	Investigation of objective measures for intelligibility prediction of noise-reduced speech for Chinese, Japanese, and English. Journal of the Acoustical Society of America, 2014, 136, 3301-3312.	1.1	5
39	Effect of the division between early and late reflections on intelligibility of ideal binary-masked speech. Journal of the Acoustical Society of America, 2015, 137, 2801-2810.	1.1	5
40	Adaptation in Mandarin tone production with pitch-shifted auditory feedback: influence of tonal contrast requirements. Language, Cognition and Neuroscience, 2018, 33, 734-749.	1.2	5
41	Self-Supervised Pre-Training for Attention-Based Encoder-Decoder ASR Model. IEEE/ACM Transactions on Audio Speech and Language Processing, 2022, 30, 1763-1774.	5.8	5
42	Using a Kind of Novel Phonotactic Information for SVM Based Speaker Recognition. IEICE Transactions on Information and Systems, 2009, E92-D, 746-749.	0.7	4
43	A Hybrid Speech Emotion Recognition System Based on Spectral and Prosodic Features. IEICE Transactions on Information and Systems, 2010, E93-D, 2813-2821.	0.7	4
44	Optimized large vocabulary WFST speech recognition system. , 2012, , .		4
45	An LVCSR Based Automatic Scoring Method in English Reading Tests. , 2012, , .		4
46	A novel discriminative method for pronunciation quality assessment. , 2013, , .		4
47	Noise Estimation Using a Constrained Sequential Hidden Markov Model in the Log-Spectral Domain. IEEE Transactions on Audio Speech and Language Processing, 2013, 21, 1145-1157.	3.2	4
48	Language recognition system using language branch discriminative information. , 2014, , .		4
49	Similar Language Identification for Uyghur and Kazakh on Short Spoken Texts. , 2016, , .		4
50	Semi-Supervised Learning with Deep Neural Networks for Relative Transfer Function Inverse Regression. , 2018, , .		4
51	Identity Vector Extraction Using Shared Mixture of PLDA for Shortâ€Time Speaker Recognition. Chinese Journal of Electronics, 2019, 28, 357-363.	1.5	4
52	A New Timeâ€“Frequency Attention Tensor Network for Language Identification. Circuits, Systems, and Signal Processing, 2020, 39, 2744-2758.	2.0	4
53	Estimation Reliability Function Assisted Sound Source Localization With Enhanced Steering Vector Phase Difference. IEEE/ACM Transactions on Audio Speech and Language Processing, 2021, 29, 421-435.	5.8	4
54	Context-dependent Label Smoothing Regularization for Attention-based End-to-End Code-Switching Speech Recognition. , 2021, , .		4

#	Article	IF	Citations
55	Spoken Term Detection Using Dynamic Match Subword Confusion Network. , 2008, , .		3
56	Acoustic Feature Optimization Based on F-Ratio for Robust Speech Recognition. IEICE Transactions on Information and Systems, 2010, E93-D, 2417-2430.	0.7	3
57	A two microphone-based approach for speech enhancement in adverse environments. , 2012, , .		3
58	Voice biometrics using linear Gaussian model. IET Biometrics, 2014, 3, 9-15.	2.5	3
59	Robust beamforming using beamâ€toâ€reference weighting diagonal loading and Bayesian framework. Electronics Letters, 2015, 51, 1772-1774.	1.0	3
60	Speech Enhancement Using Multiâ€channel Postâ€Filtering with Modified Signal Presence Probability in Reverberant Environment. Chinese Journal of Electronics, 2016, 25, 512-519.	1.5	3
61	Improvement of mask-based speech source separation using DNN. , 2016, , .		3
62	Robust speaker recognition using library of crossâ€domain variation compensation transforms. Electronics Letters, 2016, 52, 321-323.	1.0	3
63	Discriminatively learned network for iâ€vector based speaker recognition. Electronics Letters, 2018, 54, 1302-1304.	1.0	3
64	Semantic Features Based N-Best Rescoring Methods for Automatic Speech Recognition. Applied Sciences (Switzerland), 2019, 9, 5053.	2.5	3
65	History Utterance Embedding Transformer LM for Speech Recognition. , 2021, , .		3
66	Cough-based COVID-19 Detection with Multi-band Long-Short Term Memory and Convolutional Neural Networks. , 2021, , .		3
67	An individualization approach for head-related transfer function in arbitrary directions based on deep learning. JASA Express Letters, 2022, 2, .	1.1	3
68	Keyword Spotting Based on Syllable Confusion Network. , 2007, , .		2
69	Noise power estimation based on a sequential Gaussian Mixture Model. , 2011, , .		2
70	Factor analysis of Laplacian approach for speaker recognition. , 2012, , .		2
71	Noise estimation using a constrained sequential HMM IN log-spectral domain. , 2012, , .		2
72	Two-Microphone Noise Reduction Using Spatial Information-Based Spectral Amplitude Estimation. IEICE Transactions on Information and Systems, 2012, E95.D, 1454-1464.	0.7	2

#	Article	IF	Citations
73	Automatic Allophone Deriving for Korean Speech Recognition. , 2013, , .		2
74	Speeding up deep neural networks for speech recognition on ARM Cortex-A series processors. , 2014, , .		2
75	Feature recovery for noiseâ€robust speaker verification. Electronics Letters, 2015, 51, 1459-1461.	1.0	2
76	Effective utilization of multiple examples in query-by-example spoken term detection. , 2016, , .		2
77	PLF Optimization for Target Language Detection. Chinese Journal of Electronics, 2017, 26, 118-121.	1.5	2
78	Tailoring an Interpretable Neural Language Model. IEEE/ACM Transactions on Audio Speech and Language Processing, 2019, 27, 1164-1178.	5.8	2
79	A Multi-Feature Compression and Fusion Strategy of Vertical Self-Contained Hydrophone Array. IEEE Sensors Journal, 2021, 21, 24349-24358.	4.7	2
80	Far-Field Speech Recognition Based on Complex-Valued Neural Networks and Inter-Frame Similarity Difference Method. , 2021, , .		2
81	SI-Net: Multi-Scale Context-Aware Convolutional Block for Speaker Verification. , 2021, , .		2
82	A Compact-Microphone-Array-Based Speech Enhancement Algorithm Using Auditory Subbands and Probability Constrained Postfilter. , 2008, , .		1
83	Nonnative Speech Recognition Based on State-Level Bilingual Model Modification. , 2008, , .		1
84	Automatic Singing Performance Evaluation for Untrained Singers. IEICE Transactions on Information and Systems, 2009, E92-D, 1596-1600.	0.7	1
85	Development of a Chinese song name recognition system. , 2011, , .		1
86	Semantic Class Labeling Based on CRF for Limited Domain Searching Service. , 2012, , .		1
87	Factor Analysis of Neighborhood-Preserving Embedding for Speaker Verification. IEICE Transactions on Information and Systems, 2012, E95.D, 2572-2576.	0.7	1
88	Recurrent neural network language model in mandarin voice input system. , 2012, , .		1
89	Improving Korean LVCSR with Long-Time Temporal Patterns and an Extended Phoneme Set. , 2013, , .		1
90	Hybrid Reverberator Using Multiple Impulse Responses for Audio Rendering Improvement. , 2013, , .		1

#	Article	IF	Citations
91	Discriminative Approach to Build Hybrid Vocabulary for Conversational Telephone Speech Recognition of Agglutinative Languages. IEICE Transactions on Information and Systems, 2013, E96.D, 2478-2482.	0.7	1
92	Fuzzy Matching of Semantic Class in Chinese Spoken Language Understanding. IEICE Transactions on Information and Systems, 2013, E96.D, 1845-1852.	0.7	1
93	On the Performance and Robustness of Crosstalk Cancelation with Multiple Loudspeakers. , 2014, , .		1
94	A new robust auxiliary noise power scheduling for online secondary path modeling in active noise control systems. , 2014, , .		1
95	Improved mandarin spoken term detection by using deep neural network for keyword verification. , 2014, , .		1
96	Enhanced Out of Vocabulary Word Detection Using Local Acoustic Information. , 2014, , .		1
97	Reverberation robust two-microphone Target Signal Detection algorithm with coherent interference. , 2014, , .		1
98	Crossâ€domain variation compensation for robust speaker verification. Electronics Letters, 2015, 51, 1706-1707.	1.0	1
99	Structural Optimization and Online Evolutionary Learning for Spoken Dialog Management. IEEE Signal Processing Letters, 2016, 23, 1013-1017.	3.6	1
100	An unsupervised vocabulary selection technique for Chinese automatic speech recognition. , 2016, , .		1
101	Characterization Vector Extraction Using Neural Network for Speaker Recognition. , 2016, , .		1
102	Predicting user influence in microblogs. , 2016, , .		1
103	A General Bayesian Model for Speaker Verification. Chinese Journal of Electronics, 2016, 25, 1045-1051.	1.5	1
104	Dynamic group sparsity for non-negative matrix factorization with application to unsupervised source separation. , 2016, , .		1
105	A Robust Stepâ€Size Control Technique Based on Proportionate Constraints on Filter Update for Acoustic Echo Cancellation. Chinese Journal of Electronics, 2016, 25, 692-699.	1.5	1
106	Speech intelligibility enhancement in noisy reverberant conditions. , 2016, , .		1
107	Predicting user influence under the environment of big data. , 2017, , .		1
108	Collective prediction based on community structure. Physica A: Statistical Mechanics and Its Applications, 2017, 465, 587-598.	2.6	1

#	Article	IF	Citations
109	A Stochastic Approximation Method with Enhanced Robustness for Crosstalk Cancellation. Chinese Journal of Electronics, 2017, 26, 1269-1275.	1.5	1
110	Handling OOVWords in Mandarin Spoken Term Detection with an Hierarchical nâ€Gram Language Model. Chinese Journal of Electronics, 2017, 26, 1239-1244.	1.5	1
111	Singleâ€Channel Speech Separation Based on Nonâ€negative Matrix Factorization and Factorial Conditional Random Field. Chinese Journal of Electronics, 2018, 27, 1063-1070.	1.5	1
112	Language Model Score Regularization for Speech Recognition. Chinese Journal of Electronics, 2019, 28, 604-609.	1.5	1
113	Robust audio retrieval method based on antiâ€noise fingerprinting and segmental matching. Electronics Letters, 2020, 56, 245-247.	1.0	1
114	An E2E-ASR-Based Iteratively-Trained Timestamp Estimator. IEEE Signal Processing Letters, 2022, 29, 1654-1658.	3.6	1
115	State-based bilingual model modification for nonnative speech recognition. , 2008, , .		0
116	Approximate Decision Function and Optimization for GMM-UBM Based Speaker Verification. IEICE Transactions on Information and Systems, 2009, E92-D, 1798-1802.	0.7	0
117	A bayesian logistic regression approach to spoken language identification. IEICE Electronics Express, 2010, 7, 390-396.	0.8	0
118	Enhancing the Robustness of the Posterior-Based Confidence Measures Using Entropy Information for Speech Recognition. IEICE Transactions on Information and Systems, 2010, E93-D, 2431-2439.	0.7	0
119	Dual-channel optimally modified log-spectral amplitude estimator using spatial information. , 2011, , .		0
120	Quantization Index Modulation audio watermarking system using a psychoacoustic model. , 2011, , .		0
121	Robust understanding of spoken Chinese through character-based tagging and prior knowledge exploitation. , 2011, , .		0
122	ASR-Based Audio Pattern Discovery. , 2011, , .		0
123	Improved acoustic models for Conversational Telephone Speech recognition. , 2012, , .		0
124	Noise Robust Feature Scheme for Automatic Speech Recognition Based on Auditory Perceptual Mechanisms. IEICE Transactions on Information and Systems, 2012, E95.D, 1610-1618.	0.7	0
125	Logarithmic Adaptive Quantization Projection for Audio Watermarking. IEICE Transactions on Information and Systems, 2012, E95.D, 1436-1445.	0.7	0
126	Target speech detection based on microphone array using inter-channel phase differences. , 2012, , .		0

#	Article	IF	Citations
127	Impact of Word Classing on Recurrent Neural Network Language Model. , 2012, , .		0
128	An Improved Mandarin Voice Input System Using Recurrent Neural Network Language Model. , 2012, , .		0
129	Lattice generation with accurate word boundary in WFST framework. , 2012, , .		0
130	Objective Japanese intelligibility prediction for noisy speech signals before and after noise-reduction processing. , 2013, , .		0
131	Pitch estimation based on harmonic salience. Electronics Letters, 2013, 49, 1491-1492.	1.0	0
132	Speaker Recognition Using Sparse Probabilistic Linear Discriminant Analysis. IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, 2013, E96.A, 1938-1945.	0.3	0
133	Acoustic Echo Control with Frequency-Domain Stage-Wise Regression. IEEE Signal Processing Letters, 2014, 21, 1265-1269.	3.6	0
134	Smoothing Method for Improved Minimum Phone Error Linear Regression. IEICE Transactions on Information and Systems, 2014, E97.D, 2105-2113.	0.7	0
135	A Hybrid Approach for Reverberation Simulation. IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, 2015, E98.A, 2101-2108.	0.3	0
136	Discriminative Pronunciation Modeling Using the MPE Criterion. IEICE Transactions on Information and Systems, 2015, E98.D, 717-720.	0.7	0
137	Two-stage ASGD framework for parallel training of DNN acoustic models using Ethernet. , 2015, , .		0
138	Reverberation robust multi-channel post-filtering using modified signal presence probability. , 2015, , .		0
139	Equalization of Sound Reproduction System Based on the Human Perception Characteristics. , 2015, , .		0
140	Restoration of instantaneous amplitude and phase of speech signal in noisy reverberant environments. , 2015, , .		0
141	Robust multiple speech source localization based on phase difference regression. , 2016, , .		0
142	Full-posterior PLDA based speaker diarization of telephone conversations. , 2017, , .		0
143	On SDW-MWF and Variable Span Linear Filter with Application to Speech Recognition in Noisy Environments. , 2018, , .		0
144	Binaural rendering technology over loudspeakers and headphones. Acoustical Science and Technology, 2020, 41, 134-141.	0.5	0

#	ARTICLE	IF	CITATIONS
145	Improves Neural Acoustic Word Embeddings Query by Example Spoken Term Detection with Wav2vec Pretraining and Circle Loss. , 2021, , .		0
146	Using Cognitive Interest Graph and Knowledge-activated Attention for Learning Resource Recommendation. , 2021, , .		0
147	A two-element-microphone-array-based speech recognition system in vehicle environment. Acoustical Science and Technology, 2009, 30, 51-54.	0.5	0
148	Lingual-Agnostic Meta-Learning for Low-Resource Part-of-Speech Tagging. , 2020, , .		0