

Zheng-Hua Tan

List of Publications by Year in descending order

Source: <https://exaly.com/author-pdf/888561/publications.pdf>

Version: 2024-02-01

125
papers

3,027
citations

361413

20
h-index

254184

43
g-index

126
all docs

126
docs citations

126
times ranked

1811
citing authors

#	ARTICLE	IF	CITATIONS
1	Multitalker Speech Separation With Utterance-Level Permutation Invariant Training of Deep Recurrent Neural Networks. IEEE/ACM Transactions on Audio Speech and Language Processing, 2017, 25, 1901-1913.	5.8	499
2	Permutation invariant training of deep models for speaker-independent multi-talker speech separation. , 2017, , .		442
3	Speech Intelligibility Potential of General and Specialized Deep Neural Network Based Speech Enhancement Systems. IEEE/ACM Transactions on Audio Speech and Language Processing, 2017, 25, 153-167.	5.8	129
4	Conditional Generative Adversarial Networks for Speech Enhancement and Noise-Robust Speaker Verification. , 0, , .		118
5	Adaptive protection combined with machine learning for microgrids. IET Generation, Transmission and Distribution, 2019, 13, 770-779.	2.5	115
6	An Overview of Deep-Learning-Based Audio-Visual Speech Enhancement and Separation. IEEE/ACM Transactions on Audio Speech and Language Processing, 2021, 29, 1368-1396.	5.8	111
7	Decorrelation of Neutral Vector Variables: Theory and Applications. IEEE Transactions on Neural Networks and Learning Systems, 2018, 29, 129-143.	11.3	87
8	On Loss Functions for Supervised Monaural Time-Domain Speech Enhancement. IEEE/ACM Transactions on Audio Speech and Language Processing, 2020, 28, 825-838.	5.8	75
9	Low-Complexity Variable Frame Rate Analysis for Speech Recognition and Voice Activity Detection. IEEE Journal on Selected Topics in Signal Processing, 2010, 4, 798-807.	10.8	74
10	rVAD: An unsupervised segment-based robust voice activity detection method. Computer Speech and Language, 2020, 59, 1-21.	4.3	68
11	Spoofing Detection in Automatic Speaker Verification Systems Using DNN Classifiers and Dynamic Acoustic Features. IEEE Transactions on Neural Networks and Learning Systems, 2018, 29, 4633-4644.	11.3	64
12	Automatic speech recognition over error-prone wireless networks. Speech Communication, 2005, 47, 220-242.	2.8	44
13	A Joint Approach for Single-Channel Speaker Identification and Speech Separation. IEEE Transactions on Audio Speech and Language Processing, 2012, 20, 2586-2601.	3.2	37
14	Using Theatre to Study Interaction with Care Robots. International Journal of Social Robotics, 2016, 8, 457-470.	4.6	36
15	Monaural Speech Enhancement Using Deep Neural Networks by Maximizing a Short-Time Objective Intelligibility Measure. , 2018, , .		35
16	Predicting the Intelligibility of Noisy and Nonlinearly Processed Binaural Speech. IEEE/ACM Transactions on Audio Speech and Language Processing, 2016, 24, 1908-1920.	5.8	32
17	Feature selection for neutral vector in EEG signal classification. Neurocomputing, 2016, 174, 937-945.	5.9	32
18	Nonintrusive Speech Intelligibility Prediction Using Convolutional Neural Networks. IEEE/ACM Transactions on Audio Speech and Language Processing, 2018, 26, 1925-1939.	5.8	31

#	ARTICLE	IF	CITATIONS
19	Informed Sound Source Localization Using Relative Transfer Functions for Hearing Aid Applications. IEEE/ACM Transactions on Audio Speech and Language Processing, 2017, 25, 611-623.	5.8	29
20	The Sound or Silence: Investigating the Influence of Robot Noise on Proxemics. , 2018, , .		28
21	Speech enhancement using Long Short-Term Memory based recurrent Neural Networks for noise robust Speaker Verification. , 2016, , .		27
22	Deep Spoken Keyword Spotting: An Overview. IEEE Access, 2022, 10, 4169-4199.	4.2	27
23	Utterance Verification for Text-Dependent Speaker Recognition: A Comparative Assessment Using the RedDots Corpus. , 0, , .		23
24	Analysis of beamformer directed single-channel noise reduction system for hearing aid applications. , 2015, , .		22
25	Im2Sketch: Sketch generation by unconflicted perceptual grouping. Neurocomputing, 2015, 165, 338-349.	5.9	22
26	Total Variability Modeling Using Source-Specific Priors. IEEE/ACM Transactions on Audio Speech and Language Processing, 2016, 24, 504-517.	5.8	21
27	Refinement and validation of the binaural short time objective intelligibility measure for spatially diverse conditions. Speech Communication, 2018, 102, 1-13.	2.8	21
28	OSLNet: Deep Small-Sample Classification With an Orthogonal Softmax Layer. IEEE Transactions on Image Processing, 2020, 29, 6482-6495.	9.8	21
29	Exploiting Temporal Correlation of Speech for Error Robust and Bandwidth Flexible Distributed Speech Recognition. IEEE Transactions on Audio Speech and Language Processing, 2007, 15, 1391-1403.	3.2	20
30	Using Audio-Derived Affective Offset to Enhance TV Recommendation. IEEE Transactions on Multimedia, 2014, 16, 1999-2010.	7.2	20
31	Minimum Mean-Square Error Estimation of Mel-Frequency Cepstral Featuresâ€“A Theoretically Consistent Approach. IEEE/ACM Transactions on Audio Speech and Language Processing, 2015, 23, 186-197.	5.8	20
32	Text-Independent Speaker Identification Using the Histogram Transform Model. IEEE Access, 2016, 4, 9733-9739.	4.2	20
33	Developing a speaker identification system for the DARPA RATS project. , 2013, , .		19
34	Time-Contrastive Learning Based Deep Bottleneck Features for Text-Dependent Speaker Verification. IEEE/ACM Transactions on Audio Speech and Language Processing, 2019, 27, 1267-1279.	5.8	19
35	Convex Combination of Multiple Statistical Models With Application to VAD. IEEE Transactions on Audio Speech and Language Processing, 2011, 19, 2314-2327.	3.2	18
36	On the Relationship Between Short-Time Objective Intelligibility and Short-Time Spectral-Amplitude Mean-Square Error for Speech Enhancement. IEEE/ACM Transactions on Audio Speech and Language Processing, 2019, 27, 283-295.	5.8	18

#	ARTICLE	IF	CITATIONS
37	Signal-to-Signal Ratio Independent Speaker Identification for Co-channel Speech Signals. , 2010, , .		17
38	A non-intrusive Short-Time Objective Intelligibility measure. , 2017, , .		17
39	Deep-learning-based audio-visual speech enhancement in presence of Lombard effect. Speech Communication, 2019, 115, 38-50.	2.8	17
40	Personalized signal-independent beamforming for binaural hearing aids. Journal of the Acoustical Society of America, 2019, 145, 2971-2981.	1.1	17
41	Vocoder-Based Speech Synthesis from Silent Videos. , 0, , .		17
42	Noise Condition-Dependent Training Based on Noise Classification and SNR Estimation. IEEE Transactions on Audio Speech and Language Processing, 2007, 15, 2431-2443.	3.2	16
43	Adaptive overcurrent protection for microgrids in extensive distribution systems. , 2016, , .		16
44	SketchSegNet+: An End-to-End Learning of RNN for Multi-Class Sketch Semantic Segmentation. IEEE Access, 2019, 7, 102717-102726.	4.2	16
45	On Training Targets and Objective Functions for Deep-learning-based Audio-visual Speech Enhancement. , 2019, , .		14
46	Text Dependent Speaker Verification Using Un-Supervised HMM-UBM and Temporal GMM-UBM. , 0, , .		14
47	A subvector-based error concealment algorithm for speech recognition over mobile networks. , 0, , .		13
48	Joint single-channel speech separation and speaker identification. , 2010, , .		13
49	iSocioBot: A Multimodal Interactive Social Robot. International Journal of Social Robotics, 2018, 10, 5-19.	4.6	13
50	Latent Dirichlet mixture model. Neurocomputing, 2018, 278, 12-22.	5.9	13
51	Robust Speech Recognition by Nonlocal Means Denoising Processing. IEEE Signal Processing Letters, 2008, 15, 701-704.	3.6	12
52	A heuristic hierarchical scheme for academic search and retrieval. Information Processing and Management, 2013, 49, 1326-1343.	8.6	12
53	Audio-based age and gender identification to enhance the recommendation of TV content. IEEE Transactions on Consumer Electronics, 2013, 59, 721-729.	3.6	12
54	Implementing a Commercial-Strength Parallel Hybrid Movie Recommendation Engine. IEEE Intelligent Systems, 2014, 29, 92-96.	4.0	11

#	ARTICLE	IF	CITATIONS
55	Improved External Speaker-Robust Keyword Spotting for Hearing Assistive Devices. IEEE/ACM Transactions on Audio Speech and Language Processing, 2020, 28, 1233-1247.	5.8	11
56	Hybrid evolutionary approach for designing neural networks for classification. Electronics Letters, 2004, 40, 955.	1.0	10
57	Informed TDoA-based direction of arrival estimation for hearing aid applications. , 2015, , .		10
58	AMORE: design and implementation of a commercial-strength parallel hybrid movie recommendation engine. Knowledge and Information Systems, 2016, 47, 671-696.	3.2	10
59	Vocal Tract Length Perturbation for Text-Dependent Speaker Verification With Autoregressive Prediction Coding. IEEE Signal Processing Letters, 2021, 28, 364-368.	3.6	9
60	Assessing Wireless Sensing Potential With Large Intelligent Surfaces. IEEE Open Journal of the Communications Society, 2021, 2, 934-947.	6.9	9
61	A Novel Loss Function and Training Strategy for Noise-Robust Keyword Spotting. IEEE/ACM Transactions on Audio Speech and Language Processing, 2021, 29, 2254-2266.	5.8	9
62	Online Multichannel Speech Enhancement Based on Recursive EM and DNN-Based Speech Presence Estimation. IEEE/ACM Transactions on Audio Speech and Language Processing, 2020, 28, 3080-3094.	5.8	9
63	Fuzzy metagraph and its combination with the indexing approach in rule-based systems. IEEE Transactions on Knowledge and Data Engineering, 2006, 18, 829-841.	5.7	8
64	Predictive Distribution of the Dirichlet Mixture Model by Local Variational Inference. Journal of Signal Processing Systems, 2014, 74, 359-374.	2.1	8
65	Joint variable frame rate and length analysis for speech recognition under adverse conditions. Computers and Electrical Engineering, 2014, 40, 2139-2149.	4.8	8
66	Local feature learning for face recognition under varying poses. , 2015, , .		8
67	Improved Gaussian Mixture Models for Adaptive Foreground Segmentation. Wireless Personal Communications, 2016, 87, 629-643.	2.7	8
68	On the Comparisons of Decorrelation Approaches for Non-Gaussian Neutral Vector Variables. IEEE Transactions on Neural Networks and Learning Systems, 2023, 34, 1823-1837.	11.3	8
69	Network, Distributed and Embedded Speech Recognition: An Overview. Advances in Pattern Recognition, 2008, , 1-23.	0.8	8
70	Keyword Spotting for Hearing Assistive Devices Robust to External Speakers. , 0, , .		8
71	Guided spectrogram filtering for speech dereverberation. Applied Acoustics, 2018, 134, 154-159.	3.3	7
72	Incorporating pass-phrase dependent background models for text-dependent speaker verification. Computer Speech and Language, 2018, 47, 259-271.	4.3	7

#	ARTICLE	IF	CITATIONS
73	Exploring Filterbank Learning for Keyword Spotting. , 2021, , .		7
74	Partial splicing packet loss concealment for distributed speech recognition. Electronics Letters, 2003, 39, 1619.	1.0	6
75	Adaptive Multi-Frame-Rate Scheme for Distributed Speech Recognition Based on a Half Frame-Rate Front-End. , 2005, , .		6
76	Speech Recognition on Mobile Devices. Lecture Notes in Computer Science, 2010, , 221-237.	1.3	6
77	A perceptually motivated LP residual estimator in noisy and reverberant environments. Speech Communication, 2018, 96, 129-141.	2.8	6
78	Soft Dropout And Its Variational Bayes Approximation. , 2019, , .		6
79	A posteriori SNR weighted energy based variable frame rate analysis for speech recognition. , 0, , .		6
80	OOV-detection and channel error protection for distributed speech recognition over wireless networks. , 0, , .		5
81	A MAP criterion for detecting the number of speakers at frame level in model-based single-channel speech separation. , 2010, , .		5
82	Bias-Compensated Informed Sound Source Localization Using Relative Transfer Functions. IEEE/ACM Transactions on Audio Speech and Language Processing, 2018, 26, 1275-1289.	5.8	5
83	Joint Maximum Likelihood Estimation of Power Spectral Densities and Relative Acoustic Transfer Functions for Acoustic Beamforming. , 2021, , .		5
84	Designing and Implementing an Interactive Social Robot from Off-the-Shelf Components. Mechanisms and Machine Science, 2015, , 113-121.	0.5	5
85	HAPPY Team Entry to NIST OpenSAD Challenge: A Fusion of Short-Term Unsupervised and Segment i-Vector Based Speech Activity Detectors. , 0, , .		5
86	Exploitation of spectral variance to improve robustness in speech recognition. Electronics Letters, 2006, 42, 312.	1.0	4
87	EEG signal classification with super-Dirichlet mixture model. , 2012, , .		4
88	Combination of Multiple Measurement Cues for Visual Face Tracking. Wireless Personal Communications, 2014, 78, 1789-1810.	2.7	4
89	Informed Direction of Arrival estimation using a spherical-head model for Hearing Aid applications. , 2016, , .		4
90	Frame Selection for Robust Speaker Identification: A Hybrid Approach. Wireless Personal Communications, 2017, 97, 933-950.	2.7	4

#	ARTICLE	IF	CITATIONS
91	Audio-Based Granularity-Adapted Emotion Classification. IEEE Transactions on Affective Computing, 2018, 9, 176-190.	8.3	4
92	Robust Bayesian and Maximum a Posteriori Beamforming for Hearing Assistive Devices. , 2019, , .		4
93	Disentangled Speech Representation Learning Based on Factorized Hierarchical Variational Autoencoder with Self-Supervised Objective. , 2021, , .		4
94	Multichannel Speech Enhancement With Own Voice-Based Interfering Speech Suppression for Hearing Assistive Devices. IEEE/ACM Transactions on Audio Speech and Language Processing, 2022, 30, 706-720.	5.8	4
95	Using Closed-Set Speaker Identification Score Confidence to Enhance Audio-Based Collaborative Filtering for Multiple Users. IEEE Transactions on Consumer Electronics, 2018, 64, 11-18.	3.6	3
96	Effects of Lombard Reflex on the Performance of Deep-learning-based Audio-visual Speech Enhancement Systems. , 2019, , .		3
97	The Importance of Context When Recommending TV Content: Dataset and Algorithms. IEEE Transactions on Multimedia, 2020, 22, 1531-1541.	7.2	3
98	Self-segmentation of pass-phrase utterances for deep feature learning in text-dependent speaker verification. Computer Speech and Language, 2021, 70, 101229.	4.3	3
99	A Configurable Distributed Speech Recognition System. , 2007, , 59-70.		3
100	The Minimum Overlap-Gap Algorithm for Speech Enhancement. IEEE Access, 2022, 10, 14698-14716.	4.2	3
101	Guest Editorsâ€™ Introduction to the Special Issue on â€œNew Trends in Signal Processing and Biomedical Engineeringâ€• Computers and Electrical Engineering, 2012, 38, 1-2.	4.8	2
102	Maximum Likelihood Estimation of the Interference-Plus-Noise Cross Power Spectral Density Matrix for Own Voice Retrieval. , 2020, , .		2
103	UIAI System for Short-Duration Speaker Verification Challenge 2020. , 2021, , .		2
104	Mobile video annotation for enhanced rich media communication during emergency handling. , 2011, , .		2
105	Audio and Speech Processing for Data Mining. , 2009, , 98-103.		2
106	Sinusoidal approach for the single-channel speech separation and recognition challenge. , 0, , .		2
107	Demographic recommendation by means of group profile elicitation using speaker age and gender recognition. , 0, , .		2
108	Robust Speech Recognition From Noise-Type Based Feature Compensation and Model Interpolation in a Multiple Model Framework. , 0, , .		1

#	ARTICLE	IF	CITATIONS
109	<title>Variable frame rate analysis for automatic speech recognition</title>. Proceedings of SPIE, 2007, , .	0.8	1
110	Perceptual grouping via untangling Gestalt principles. , 2013, , .		1
111	Exploring super-gaussianity toward robust information-theoretical time delay estimation. Journal of the Acoustical Society of America, 2013, 133, 1515-1524.	1.1	1
112	Visual Detection of Events of Interest from Urban Activity. Wireless Personal Communications, 2017, 97, 1877-1888.	2.7	1
113	Training Data-Driven Speech Intelligibility Predictors on Heterogeneous Listening Test Data. IEEE Access, 2022, 10, 66175-66189.	4.2	1
114	Introduction to the Issue on Speech Processing for Natural Interaction With Intelligent Environments. IEEE Journal on Selected Topics in Signal Processing, 2010, 4, 769-771.	10.8	0
115	Projecting emotional speech into arousal-valence space using pairwise preference learning. , 2016, , .		0
116	Dirichlet mixture allocation. , 2016, , .		0
117	Background subtraction for patterns of activities in cities. , 2016, , .		0
118	Highlights From the Machine Learning for Signal Processing Technical Committee [In the Spotlight]. IEEE Signal Processing Magazine, 2020, 37, 202-200.	5.6	0
119	Can robots express facial emotions dominantly enough for use in dementia care?. International Psychogeriatrics, 2020, 32, 891-892.	1.0	0
120	High-accuracy, low-complexity voice activity detection based on a posteriori SNR weighted energy. , 0, , .		0
121	A system for detecting miscues in dyslexic read speech. , 0, , .		0
122	Multi-sensor voice activity detection based on multiple observation hypothesis testing. , 0, , .		0
123	Utilising Tree-Based Ensemble Learning for Speaker Segmentation. Lecture Notes in Computer Science, 2014, , 50-59.	1.3	0
124	Active Reconfigurable Operation with Long Short-Term Memory Prediction for Smart City Microgrids. , 2019, , .		0
125	iVAE-GAN: Identifiable VAE-GAN Models for Latent Representation Learning. IEEE Access, 2022, 10, 48405-48418.	4.2	0