

Shrikanth S Narayanan

List of Publications by Year in descending order

Source: <https://exaly.com/author-pdf/2899100/publications.pdf>

Version: 2024-02-01

408
papers

12,437
citations

66343

42
h-index

46799

89
g-index

422
all docs

422
docs citations

422
times ranked

7060
citing authors

#	ARTICLE	IF	CITATIONS
1	IEMOCAP: interactive emotional dyadic motion capture database. <i>Language Resources and Evaluation</i> , 2008, 42, 335-359.	2.7	1,756
2	Acoustics of children's speech: Developmental changes of temporal and spectral parameters. <i>Journal of the Acoustical Society of America</i> , 1999, 105, 1455-1468.	1.1	609
3	Environmental Sound Recognition With Time-Frequency Audio Features. <i>IEEE Transactions on Audio Speech and Language Processing</i> , 2009, 17, 1142-1158.	3.2	475
4	Primitives-based evaluation and estimation of emotions in speech. <i>Speech Communication</i> , 2007, 49, 787-800.	2.8	308
5	Emotion recognition using a hierarchical binary decision tree approach. <i>Speech Communication</i> , 2011, 53, 1162-1171.	2.8	274
6	An approach to real-time magnetic resonance imaging for speech production. <i>Journal of the Acoustical Society of America</i> , 2004, 115, 1771-1776.	1.1	256
7	The Vera am Mittag German audio-visual emotional speech database. , 2008, , .		246
8	Behavioral Signal Processing: Deriving Human Behavioral Informatics From Speech and Language. <i>Proceedings of the IEEE</i> , 2013, 101, 1203-1233.	21.3	225
9	Analysis of Emotionally Salient Aspects of Fundamental Frequency for Emotion Detection. <i>IEEE Transactions on Audio Speech and Language Processing</i> , 2009, 17, 582-596.	3.2	211
10	Paralinguistics in speech and language—State-of-the-art and the challenge. <i>Computer Speech and Language</i> , 2013, 27, 4-39.	4.3	207
11	Toward articulatory-acoustic models for liquid approximants based on MRI and EPG data. Part I. The laterals. <i>Journal of the Acoustical Society of America</i> , 1997, 101, 1064-1077.	1.1	162
12	Robust Voice Activity Detection Using Long-Term Signal Variability. <i>IEEE Transactions on Audio Speech and Language Processing</i> , 2011, 19, 600-613.	3.2	153
13	Applying Machine Learning to Facilitate Autism Diagnostics: Pitfalls and Promises. <i>Journal of Autism and Developmental Disorders</i> , 2015, 45, 1121-1136.	2.7	149
14	Rigid Head Motion in Expressive Speech Animation: Analysis and Synthesis. <i>IEEE Transactions on Audio Speech and Language Processing</i> , 2007, 15, 1075-1086.	3.2	147
15	Use of machine learning to improve autism screening and diagnostic instruments: effectiveness, efficiency, and multi-instrument fusion. <i>Journal of Child Psychology and Psychiatry and Allied Disciplines</i> , 2016, 57, 927-937.	5.2	140
16	Automatic speaker age and gender recognition using acoustic and prosodic level information fusion. <i>Computer Speech and Language</i> , 2013, 27, 151-167.	4.3	135
17	Context-Sensitive Learning for Enhanced Audiovisual Emotion Classification. <i>IEEE Transactions on Affective Computing</i> , 2012, 3, 184-198.	8.3	131
18	MUPET—Mouse Ultrasonic Profile ExTraction: A Signal Processing Tool for Rapid and Unsupervised Analysis of Ultrasonic Vocalizations. <i>Neuron</i> , 2017, 94, 465-485.e5.	8.1	131

#	ARTICLE	IF	CITATIONS
19	Optimal Arousal Identification and Classification for Affective Computing Using Physiological Signals: Virtual Reality Stroop Task. IEEE Transactions on Affective Computing, 2010, 1, 109-118.	8.3	120
20	Real-time magnetic resonance imaging and electromagnetic articulography database for speech production research (TC). Journal of the Acoustical Society of America, 2014, 136, 1307-1311.	1.1	120
21	Synchronized and noise-robust audio recordings during realtime magnetic resonance imaging scans. Journal of the Acoustical Society of America, 2006, 120, 1791-1794.	1.1	104
22	Interrelation Between Speech and Facial Gestures in Emotional Utterances: A Single Subject Study. IEEE Transactions on Audio Speech and Language Processing, 2007, 15, 2331-2347.	3.2	101
23	Toward articulatory-acoustic models for liquid approximants based on MRI and EPG data. Part II. The rhotics. Journal of the Acoustical Society of America, 1997, 101, 1078-1089.	1.1	100
24	Annotation and processing of continuous emotional attributes: Challenges and opportunities. , 2013, , .		99
25	A review of speaker diarization: Recent advances with deep learning. Computer Speech and Language, 2022, 72, 101317.	4.3	90
26	Interpreting ambiguous emotional expressions. , 2009, , .		89
27	Region Segmentation in the Frequency Domain Applied to Upper Airway Real-Time Magnetic Resonance Images. IEEE Transactions on Medical Imaging, 2009, 28, 323-338.	8.9	87
28	Tracking continuous emotional trends of participants during affective dyadic interactions using body language and speech information. Image and Vision Computing, 2013, 31, 137-152.	4.5	87
29	Toward automating a human behavioral coding system for married couples'™ interactions using speech acoustic features. Speech Communication, 2013, 55, 1-21.	2.8	84
30	Seeing speech: Capturing vocal tract shaping using real-time magnetic resonance imaging [Exploratory DSP]. IEEE Signal Processing Magazine, 2008, 25, 123-132.	5.6	82
31	Signal Processing and Machine Learning for Mental Health Research and Clinical Applications [Perspectives]. IEEE Signal Processing Magazine, 2017, 34, 196-195.	5.6	73
32	Data Augmentation Using GANs for Speech Emotion Recognition. , 0, , .		73
33	Detecting emotional state of a child in a conversational computer game. Computer Speech and Language, 2011, 25, 29-44.	4.3	72
34	"Rate My Therapist": Automated Detection of Empathy in Drug and Alcohol Counseling via Speech and Language Processing. PLoS ONE, 2015, 10, e0143055.	2.5	72
35	Support Vector Regression for Automatic Recognition of Spontaneous Emotions in Speech. , 2007, , .		68
36	Robust ECG Biometrics by Fusing Temporal and Cepstral Information. , 2010, , .		68

#	ARTICLE	IF	CITATIONS
37	Timing effects of syllable structure and stress on nasals: A real-time MRI examination. Journal of Phonetics, 2009, 37, 97-110.	1.2	64
38	Decision level combination of multiple modalities for recognition and analysis of emotional expression. , 2010, , .		64
39	Audio-Visual Emotion Recognition Using Gaussian Mixture Models for Face and Voice. , 2008, , .		63
40	Accelerated three-dimensional upper airway MRI using compressed sensing. Magnetic Resonance in Medicine, 2009, 61, 1434-1440.	3.0	63
41	Multimodal Prediction of Affective Dimensions and Depression in Human-Computer Interactions. , 2014, , .		63
42	Real-time Emotion Detection System using Speech: Multi-modal Fusion of Different Timescale Features. , 2007, , .		61
43	Theoretical Analysis of Diversity in an Ensemble of Automatic Speech Recognition Systems. IEEE/ACM Transactions on Audio Speech and Language Processing, 2014, 22, 711-726.	5.8	59
44	Automatic intelligibility classification of sentence-level pathological speech. Computer Speech and Language, 2015, 29, 132-144.	4.3	59
45	Distributional Semantic Models for Affective Text Analysis. IEEE Transactions on Audio Speech and Language Processing, 2013, 21, 2379-2392.	3.2	58
46	A generalized smoothness criterion for acoustic-to-articulatory inversion. Journal of the Acoustical Society of America, 2010, 128, 2162-2172.	1.1	56
47	Designing Contestability. , 2017, 2017, 95-99.		56
48	Geometry, kinematics, and acoustics of Tamil liquid consonants. Journal of the Acoustical Society of America, 1999, 106, 1993-2007.	1.1	54
49	Speech emotion estimation in 3D space. , 2010, , .		53
50	A fast and flexible MRI system for the study of dynamic vocal tract shaping. Magnetic Resonance in Medicine, 2017, 77, 112-125.	3.0	53
51	Robust Unsupervised Arousal Rating:A Rule-Based Framework with Knowledge-Inspired Vocal Features. IEEE Transactions on Affective Computing, 2014, 5, 201-213.	8.3	52
52	A Computational Study of Expressive Facial Dynamics in Children with Autism. IEEE Transactions on Affective Computing, 2018, 9, 14-20.	8.3	51
53	An individually tailored family-centered intervention for pediatric obesity in primary care: study protocol of a randomized type II hybrid effectiveness implementation trial (Raising Healthy Children) Tj ETQq1 1 0.978431419 BT /Over	10.7	49
54	Rachel: Design of an emotionally targeted interactive agent for children with autism. , 2011, , .		48

#	ARTICLE	IF	CITATIONS
55	Auto-Tuning Spectral Clustering for Speaker Diarization Using Normalized Maximum Eigengap. IEEE Signal Processing Letters, 2020, 27, 381-385.	3.6	48
56	Flexible retrospective selection of temporal resolution in real-time speech MRI using a golden-ratio spiral view order. Magnetic Resonance in Medicine, 2011, 65, 1365-1371.	3.0	47
57	Design feasibility of an automated, machine-learning based feedback system for motivational interviewing.. Psychotherapy, 2019, 56, 318-328.	1.2	47
58	Iterative Feature Normalization Scheme for Automatic Emotion Detection from Speech. IEEE Transactions on Affective Computing, 2013, 4, 386-397.	8.3	46
59	Pathological Voice Assessment. , 2006, 2006, 1669-73.		45
60	Environmental sound recognition using MP-based features. Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, 2008, , .	1.8	44
61	On Short-Time Estimation of Vocal Tract Length from Formant Frequencies. PLoS ONE, 2015, 10, e0132193.	2.5	44
62	Clinical state tracking in serious mental illness through computational analysis of speech. PLoS ONE, 2020, 15, e0225695.	2.5	44
63	Strategies to Improve the Robustness of Agglomerative Hierarchical Clustering Under Data Source Variation for Speaker Diarization. IEEE Transactions on Audio Speech and Language Processing, 2008, 16, 1590-1601.	3.2	43
64	Sparse Representation of Electrodermal Activity With Knowledge-Driven Dictionaries. IEEE Transactions on Biomedical Engineering, 2015, 62, 960-971.	4.2	42
65	Irregularity-Aware Graph Fourier Transforms. IEEE Transactions on Signal Processing, 2018, 66, 5746-5761.	5.3	42
66	Simplified supervised i-vector modeling with application to robust and efficient language identification and speaker verification. Computer Speech and Language, 2014, 28, 940-958.	4.3	41
67	The USC CreativeIT database of multimodal dyadic interactions: from speech and full body motion capture to continuous emotional annotations. Language Resources and Evaluation, 2016, 50, 497-521.	2.7	41
68	Visual emotion recognition using compact facial representations and viseme information. , 2010, , .		40
69	Iterative feature normalization for emotional speech detection. , 2011, , .		40
70	Emotion recognition using a hierarchical binary decision tree approach. , 0, , .		37
71	Automatic Prediction of Children's Reading Ability for High-Level Literacy Assessment. IEEE Transactions on Audio Speech and Language Processing, 2011, 19, 1015-1028.	3.2	36
72	Head Motion Modeling for Human Behavior Analysis in Dyadic Interaction. IEEE Transactions on Multimedia, 2015, 17, 1107-1119.	7.2	36

#	ARTICLE	IF	CITATIONS
73	Analysis of speech production real-time MRI. <i>Computer Speech and Language</i> , 2018, 52, 1-22.	4.3	36
74	TILES-2018, a longitudinal physiologic and behavioral data set of hospital workers. <i>Scientific Data</i> , 2020, 7, 354.	5.3	36
75	Interspeaker Variability in Hard Palate Morphology and Vowel Production. <i>Journal of Speech, Language, and Hearing Research</i> , 2013, 56, 1924-1933.	1.6	35
76	Adversarial attack and defense strategies for deep speaker recognition systems. <i>Computer Speech and Language</i> , 2021, 68, 101199.	4.3	35
77	Morphological Variation in the Adult Hard Palate and Posterior Pharyngeal Wall. <i>Journal of Speech, Language, and Hearing Research</i> , 2013, 56, 521-530.	1.6	34
78	Automatic speech recognition using articulatory features from subject-independent acoustic-to-articulatory inversion. <i>Journal of the Acoustical Society of America</i> , 2011, 130, EL251-EL257.	1.1	33
79	Coregulation of therapist and client emotion during psychotherapy. <i>Psychotherapy Research</i> , 2020, 30, 591-603.	1.8	33
80	“œt sounds like...“ A natural language processing approach to detecting counselor reflections in motivational interviewing.. <i>Journal of Counseling Psychology</i> , 2016, 63, 343-350.	2.0	33
81	A technology prototype system for rating therapist empathy from audio recordings in addiction counseling. <i>PeerJ Computer Science</i> , 2016, 2, e59.	4.5	32
82	Dynamic 3-D Visualization of Vocal Tract Shaping During Speech. <i>IEEE Transactions on Medical Imaging</i> , 2013, 32, 838-848.	8.9	31
83	Investigating Implicit Cues for User State Estimation in Human-Robot Interaction Using Physiological Measurements. , 2007, , .		30
84	Combining lexical, syntactic and prosodic cues for improved online dialog act tagging. <i>Computer Speech and Language</i> , 2009, 23, 407-422.	4.3	30
85	New Frontiers in Ambulatory Assessment. <i>Social Psychological and Personality Science</i> , 2017, 8, 552-563.	3.9	30
86	Behaviorally-based couple therapies reduce emotional arousal during couple conflict. <i>Behaviour Research and Therapy</i> , 2015, 72, 49-55.	3.1	29
87	Automated evaluation of psychotherapy skills using speech and language technologies. <i>Behavior Research Methods</i> , 2022, 54, 690-711.	4.0	28
88	Automatic recognition of emotion evoked by general sound events. , 2012, , .		27
89	Improved imaging of lingual articulation using real-time multislice MRI. <i>Journal of Magnetic Resonance Imaging</i> , 2012, 35, 943-948.	3.4	27
90	Computational Analysis and Simulation of Empathic Behaviors: a Survey of Empathy Modeling with Behavioral Signal Processing Framework. <i>Current Psychiatry Reports</i> , 2016, 18, 49.	4.5	27

#	ARTICLE	IF	CITATIONS
91	“That’s Aggravating, Very Aggravating” Is It Possible to Classify Behaviors in Couple Interactions Using Automatically Derived Lexical Features?. Lecture Notes in Computer Science, 2011, , 87-96.	1.3	27
92	A System for Technology Based Assessment of Language and Literacy in Young Children: the Role of Multiple Information Sources. , 2007, , .		26
93	A Globally-Variant Locally-Constant Model for Fusion of Labels from Multiple Diverse Experts without Using Reference Labels. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35, 769-783.	13.9	26
94	3D dynamic MRI of the vocal tract during natural speech. Magnetic Resonance in Medicine, 2019, 81, 1511-1520.	3.0	26
95	Modeling therapist empathy through prosody in drug addiction counseling. , 0, , .		26
96	Audio retrieval by latent perceptual indexing. Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, 2008, , .	1.8	25
97	Evaluation of Swallow Function After Tongue Cancer Treatment Using Real-Time Magnetic Resonance Imaging. JAMA Otolaryngology - Head and Neck Surgery, 2013, 139, 1312.	2.2	25
98	Multimodal Human and Environmental Sensing for Longitudinal Behavioral Studies in Naturalistic Settings: Framework for Sensor Selection, Deployment, and Management. Journal of Medical Internet Research, 2019, 21, e12832.	4.3	25
99	Predicting therapist empathy in motivational interviews using language features inspired by psycholinguistic norms. , 0, , .		25
100	On quantifying facial expression-related atypicality of children with Autism Spectrum Disorder. , 2015, 2015, 803-807.		24
101	Dynamic offâ€resonance correction for spiral realâ€time MRI of speech. Magnetic Resonance in Medicine, 2019, 81, 234-246.	3.0	24
102	Deblurring for spiral realâ€time MRI using convolutional neural networks. Magnetic Resonance in Medicine, 2020, 84, 3438-3452.	3.0	24
103	Speaker-Invariant Affective Representation Learning via Adversarial Training. , 2020, , .		24
104	Behavioral Coding of Therapist Language in Addiction Counseling Using Recurrent Neural Networks. , 0, , .		24
105	A Deep Learning Approach to Modeling Empathy in Addiction Counseling. , 0, , .		24
106	Modeling therapist empathy and vocal entrainment in drug addiction counseling. , 0, , .		24
107	TILES audio recorder. , 2018, , .		23
108	TBALL data collection: the making of a young children's speech corpus. , 0, , .		23

#	ARTICLE	IF	CITATIONS
109	Analyzing speech rate entrainment and its relation to therapist empathy in drug addiction counseling. , 0, , .		23
110	Quantifying atypicality in affective facial expressions of children with autism spectrum disorders. , 2013, 2013, 1-6.		22
111	Pathological speech processing: State-of-the-art, current challenges, and future directions. , 2016, , .		22
112	Speaker verification based on the fusion of speech acoustics and inverted articulatory signals. Computer Speech and Language, 2016, 36, 196-211.	4.3	22
113	Modeling Dynamics of Expressive Body Gestures In Dyadic Interactions. IEEE Transactions on Affective Computing, 2017, 8, 369-381.	8.3	22
114	Predicting couple therapy outcomes based on speech acoustic features. PLoS ONE, 2017, 12, e0185123.	2.5	22
115	A multimodal analysis of physical activity, sleep, and work shift in nurses with wearable sensor data. Scientific Reports, 2021, 11, 8693.	3.3	22
116	Object classification in sidescan sonar images with sparse representation techniques. , 2012, , .		21
117	Lessons Learned: Recommendations For Implementing a Longitudinal Study Using Wearable and Environmental Sensors in a Health Care Organization. JMIR MHealth and UHealth, 2019, 7, e13305.	3.7	21
118	Advancing Methods for Reliably Assessing Motivational Interviewing Fidelity using the Motivational Interviewing Skills Code. Journal of Substance Abuse Treatment, 2015, 49, 50-57.	2.8	20
119	Characterizing Articulation in Apraxic Speech Using Real-Time Magnetic Resonance Imaging. Journal of Speech, Language, and Hearing Research, 2017, 60, 877-891.	1.6	20
120	Semantic Edge Detection for Tracking Vocal Tract Air-Tissue Boundaries in Real-Time Magnetic Resonance Images. , 0, , .		20
121	Improved Speech Recognition using Acoustic and Lexical Correlates of Pitch Accent in a N-Best Rescoring Framework. , 2007, , .		19
122	Enhanced Sparse Imputation Techniques for a Robust Speech Recognition Front-End. IEEE Transactions on Audio Speech and Language Processing, 2011, 19, 2418-2429.	3.2	19
123	Real-Time Monitoring of Participants' Interaction in a Meeting using Audio-Visual Sensors. , 2007, , .		18
124	A Generative Student Model for Scoring Word Reading Skills. IEEE Transactions on Audio Speech and Language Processing, 2011, 19, 348-360.	3.2	18
125	Novel Variations of Group Sparse Regularization Techniques With Applications to Noise Robust Automatic Speech Recognition. IEEE Transactions on Audio Speech and Language Processing, 2012, 20, 1337-1346.	3.2	18
126	Intoxicated speech detection: A fusion framework with speaker-normalized hierarchical functionals and GMM supervectors. Computer Speech and Language, 2014, 28, 375-391.	4.3	18

#	ARTICLE	IF	CITATIONS
127	Advances in real-time magnetic resonance imaging of the vocal tract for speech science and technology research. APSIPA Transactions on Signal and Information Processing, 2016, 5, .	3.3	18
128	A multimodal mixture-of-experts model for dynamic emotion prediction in movies. , 2016, , .		18
129	Semi-Supervised and Transfer Learning Approaches for Low Resource Sentiment Classification. , 2018, , .		18
130	Motion-Capture Patterns of Voluntarily Mimicked Dynamic Facial Expressions in Children and Adolescents With and Without ASD. Journal of Autism and Developmental Disorders, 2019, 49, 1062-1079.	2.7	18
131	Intrapersonal and interpersonal vocal affect dynamics during psychotherapy.. Journal of Consulting and Clinical Psychology, 2021, 89, 227-239.	2.0	18
132	Computational Media Intelligence: Human-Centered Machine Analysis of Media. Proceedings of the IEEE, 2021, 109, 891-910.	21.3	18
133	A dialog act tagging approach to behavioral coding: a case study of addiction counseling conversations. , 0, , .		18
134	A hierarchical static-dynamic framework for emotion classification. , 2011, , .		17
135	Novel 16-channel receive coil array for accelerated upper airway MRI at 3 Tesla. Magnetic Resonance in Medicine, 2011, 65, 1711-1717.	3.0	17
136	Speech in Affective Computing. , 2015, , .		17
137	Rapid Language Identification. IEEE/ACM Transactions on Audio Speech and Language Processing, 2015, 23, 1118-1129.	5.8	17
138	Feasibility of through-time spiral generalized autocalibrating partial parallel acquisition for low latency accelerated real-time MRI of speech. Magnetic Resonance in Medicine, 2017, 78, 2275-2282.	3.0	17
139	Multi-Label Multi-Task Deep Learning for Behavioral Coding. IEEE Transactions on Affective Computing, 2022, 13, 508-518.	8.3	17
140	Neural Speech Decoding During Audition, Imagination and Production. IEEE Access, 2020, 8, 149714-149729.	4.2	17
141	Efficient scalable encoding for distributed speech recognition. Speech Communication, 2006, 48, 888-902.	2.8	16
142	A Statistical Approach for Modeling Prosody Features using POS Tags for Emotional Speech Synthesis. , 2007, , .		16
143	Predicting interruptions in dyadic spoken interactions. , 2010, , .		16
144	Accurate transcription of broadcast news speech using multiple noisy transcribers and unsupervised reliability metrics. , 2011, , .		16

#	ARTICLE	IF	CITATIONS
145	Long-Term SNR Estimation of Speech Signals in Known and Unknown Channel Conditions. IEEE/ACM Transactions on Audio Speech and Language Processing, 2016, 24, 2495-2506.	5.8	16
146	Test-retest repeatability of human speech biomarkers from static and real-time dynamic magnetic resonance imaging. Journal of the Acoustical Society of America, 2017, 141, 3323-3336.	1.1	16
147	A multispeaker dataset of raw and reconstructed speech production real-time MRI video and 3D volumetric images. Scientific Data, 2021, 8, 187.	5.3	16
148	Adaptive speaker identification with audiovisual cues for movie content analysis. Pattern Recognition Letters, 2004, 25, 777-791.	4.2	15
149	A review of the acoustic and linguistic properties of children's speech. , 2007, , .		15
150	Analysis of interaction attitudes using data-driven hand gesture phrases. , 2014, , .		15
151	Analysis of engagement behavior in children during dyadic interactions using prosodic cues. Computer Speech and Language, 2016, 37, 47-66.	4.3	15
152	Explaining Coronal Reduction: Prosodic Structure and Articulatory Posture. Phonetica, 2018, 75, 151-181.	0.6	15
153	Vocal tract shaping of emotional speech. Computer Speech and Language, 2020, 64, 101100.	4.3	15
154	Multimodal Embeddings From Language Models for Emotion Recognition in the Wild. IEEE Signal Processing Letters, 2021, 28, 608-612.	3.6	15
155	State-of-the-Art MRI Protocol for Comprehensive Assessment of Vocal Tract Structure and Function. , 0, , .		15
156	Pronunciation variations of Spanish-accented English spoken by young children. , 0, , .		15
157	Joint Analysis of the Emotional Fingerprint in the Face and Speech: A single subject study. , 2007, , .		14
158	Automatic Detection of Disfluency Boundaries in Spontaneous Speech of Children Using Audio–Visual Information. IEEE Transactions on Audio Speech and Language Processing, 2009, 17, 2-12.	3.2	14
159	Speaker verification using simplified and supervised i-vector modeling. , 2013, , .		14
160	A kinematic study of critical and non-critical articulators in emotional speech production. Journal of the Acoustical Society of America, 2015, 137, 1411-1429.	1.1	14
161	Modeling Multiple Time Series Annotations as Noisy Distortions of the Ground Truth: An Expectation-Maximization Approach. IEEE Transactions on Affective Computing, 2018, 9, 76-89.	8.3	14
162	Violence Rating Prediction from Movie Scripts. Proceedings of the AAAI Conference on Artificial Intelligence, 2019, 33, 671-678.	4.9	14

#	ARTICLE	IF	CITATIONS
163	Imputing Missing Data In Large-Scale Multivariate Biomedical Wearable Recordings Using Bidirectional Recurrent Neural Networks With Temporal Activation Regularization. , 2019, 2019, 2529-2534.		14
164	Robust Speaker Recognition Using Unsupervised Adversarial Invariance. , 2020, , .		14
165	Database of Volumetric and Real-Time Vocal Tract MRI for Speech Science. , 0, , .		14
166	Language Features for Automated Evaluation of Cognitive Behavior Psychotherapy Sessions. , 0, , .		14
167	Assessment of emerging reading skills in young native speakers and language learners. Speech Communication, 2009, 51, 968-984.	2.8	13
168	Processing speech signal using auditory-like filterbank provides least uncertainty about articulatory gestures. Journal of the Acoustical Society of America, 2011, 129, 4014-4022.	1.1	13
169	Speaker states recognition using latent factor analysis based Eigenchannel factor vector modeling. , 2012, , .		13
170	Developmental acoustic study of American English diphthongs. Journal of the Acoustical Society of America, 2014, 136, 1880-1894.	1.1	13
171	Online Affect Tracking with Multimodal Kalman Filters. , 2016, , .		13
172	Strategies for Disseminating Information on Biomedical Research on Autism to Hispanic Parents. Journal of Autism and Developmental Disorders, 2016, 46, 1038-1050.	2.7	13
173	Unsupervised Discovery of Character Dictionaries in Animation Movies. IEEE Transactions on Multimedia, 2018, 20, 539-551.	7.2	13
174	Family-of-origin aggression, dating aggression, and physiological stress reactivity in daily life. Physiology and Behavior, 2019, 206, 85-92.	2.1	13
175	A Bayesian network classifier for word-level reading assessment. , 0, , .		13
176	Speaker Diarization with Lexical Information. , 0, , .		13
177	Automated quality assessment of cognitive behavioral therapy sessions through highly contextualized language representations. PLoS ONE, 2021, 16, e0258639.	2.5	13
178	Universal Consistency of Data-Driven Partitions for Divergence Estimation. , 2007, , .		12
179	An exploratory study of manifolds of emotional speech. , 2010, , .		12
180	Analyzing the memory of BLSTM Neural Networks for enhanced emotion classification in dyadic spoken interactions. , 2012, , .		12

#	ARTICLE	IF	CITATIONS
181	Using physiology and language cues for modeling verbal response latencies of children with ASD. , 2013, , .		12
182	Co-registration of speech production datasets from electromagnetic articulography and real-time magnetic resonance imaging. Journal of the Acoustical Society of America, 2014, 135, EL115-EL121.	1.1	12
183	A supervised signal-to-noise ratio estimation of speech signals. , 2014, , .		12
184	Computationally deconstructing movie narratives: An informatics approach. , 2015, , .		12
185	Multimodal detection of fake social media use through a fusion of classification and pairwise ranking systems. , 2017, , .		12
186	Speed-accuracy tradeoffs in human speech production. PLoS ONE, 2018, 13, e0202180.	2.5	12
187	On Evaluating CNN Representations for Low Resource Medical Image Classification. , 2019, , .		12
188	Task-dependence of articulator synergies. Journal of the Acoustical Society of America, 2019, 145, 1504-1520.	1.1	12
189	Health behaviour outcomes of a family based intervention for paediatric obesity in primary care: A randomized type <sc>ll</sc> hybrid effectivenessâ€”implementation trial. Pediatric Obesity, 2021, 16, e12780.	2.8	12
190	Romantic partner presence and physiological responses in daily life: Attachment style as a moderator. Biological Psychology, 2021, 161, 108082.	2.2	12
191	Linguistic analysis of differences in portrayal of movie characters. , 2017, , .		12
192	Pitch Contour Stylization Using an Optimal Piecewise Polynomial Approximation. IEEE Signal Processing Letters, 2009, 16, 810-813.	3.6	11
193	Saliency-driven unstructured acoustic scene classification using latent perceptual indexing. , 2009, , .		11
194	Estimation of ordinal approach-avoidance labels in dyadic interactions: Ordinal logistic regression approach. , 2011, , .		11
195	Gender Representation in Cinematic Content. , 2015, , .		11
196	Detecting paralinguistic events in audio stream using context in features and probabilistic decisions. Computer Speech and Language, 2016, 36, 72-92.	4.3	11
197	A Novel Method for Human Bias Correction of Continuous- Time Annotations. , 2018, , .		11
198	Toward Visual Voice Activity Detection for Unconstrained Videos. , 2019, , .		11

#	ARTICLE	IF	CITATIONS
199	A modular architecture for articulatory synthesis from gestural specification. Journal of the Acoustical Society of America, 2019, 146, 4458-4471.	1.1	11
200	Automatic Prediction of Suicidal Risk in Military Couples Using Multimodal Interaction Cues from Couples Conversations. , 2020, , .		11
201	Meta-Learning With Latent Space Clustering in Generative Adversarial Network for Speaker Diarization. IEEE/ACM Transactions on Audio Speech and Language Processing, 2021, 29, 1204-1219.	5.8	11
202	Improved 3D real-time MRI of speech production. Magnetic Resonance in Medicine, 2021, 85, 3182-3195.	3.0	11
203	A Multi-task Approach to Learning Multilingual Representations. , 2018, , .		11
204	Using Prosodic and Lexical Information for Learning Utterance-level Behaviors in Psychotherapy. , 2018, 2018, 3413-3417.		11
205	A Multimodal View into Music's Effect on Human Neural, Physiological, and Emotional Experience. , 2019, , .		10
206	Robust Speech Activity Detection in Movie Audio: Data Resources and Experimental Evaluation. , 2019, , .		10
207	Improving Gender Identification in Movie Audio Using Cross-Domain Data. , 0, , .		10
208	Exploiting prosodic features for dialog act tagging in a discriminative modeling framework. , 0, , .		10
209	Robust speaker clustering strategies to data source variation for improved speaker diarization. , 2007, , .		9
210	Robust talking face video verification using joint factor analysis and sparse representation on GMM mean shifted supervectors. , 2011, , .		9
211	Are Articulatory Settings Mechanically Advantageous for Speech Motor Control?. PLoS ONE, 2014, 9, e104168.	2.5	9
212	Acoustic Denoising Using Dictionary Learning With Spectral and Temporal Regularization. IEEE/ACM Transactions on Audio Speech and Language Processing, 2018, 26, 967-980.	5.8	9
213	Speaker Agnostic Foreground Speech Detection from Audio Recordings in Workplace Settings from Wearable Recorders. , 2019, , .		9
214	Intermittently tagged real-time MRI reveals internal tongue motion during speech production. Magnetic Resonance in Medicine, 2019, 82, 600-613.	3.0	9
215	Estimating Individualized Daily Self-Reported Affect with Wearable Sensors. , 2019, , .		9
216	A computational lens into how music characterizes genre in film. PLoS ONE, 2021, 16, e0249957.	2.5	9

#	ARTICLE	IF	CITATIONS
217	An Expectation Maximization Approach to Joint Modeling of Multidimensional Ratings Derived from Multiple Annotators. , 0, , .		9
218	Convex Hull Convolutional Non-Negative Matrix Factorization for Uncovering Temporal Patterns in Multivariate Time-Series Data. , 0, , .		9
219	Leveraging Social Networks for the Assessment and Management of Neurological Patients. Seminars in Neurology, 2022, 42, 136-148.	1.4	9
220	Statistical Modeling and Retrieval of Polyphonic Music. , 2007, , .		8
221	Novel inter-cluster distance measure combining GLR and ICR for improved agglomerative hierarchical speaker clustering. Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, 2008, , .	1.8	8
222	Directional descriptors using zernike moment phases for object orientation estimation in underwater sonar images. , 2011, , .		8
223	An acoustic analysis of shared enjoyment in ECA interactions of children with autism. , 2012, , .		8
224	Data driven modeling of head motion towards analysis of behaviors in couple interactions. , 2013, , .		8
225	Markov Chain Monte Carlo Inference of Parametric Dictionaries for Sparse Bayesian Approximations. IEEE Transactions on Signal Processing, 2016, 64, 3077-3092.	5.3	8
226	Estimation of vocal tract area function from volumetric Magnetic Resonance Imaging. , 2017, , .		8
227	An Evaluation of EEG-based Metrics for Engagement Assessment of Distance Learners. , 2018, 2018, 307-310.		8
228	Toward Robust Interpretable Human Movement Pattern Analysis in a Workplace Setting. , 2019, , .		8
229	Speaker Diarization Using Latent Space Clustering in Generative Adversarial Network. , 2020, , .		8
230	Evidence of Task-Independent Person-Specific Signatures in EEG Using Subspace Techniques. IEEE Transactions on Information Forensics and Security, 2021, 16, 2856-2871.	6.9	8
231	An Affect Prediction Approach Through Depression Severity Parameter Incorporation in Neural Networks. , 0, , .		8
232	Modeling Interpersonal Linguistic Coordination in Conversations Using Word Moverâ€™s Distance. , 0, , .		8
233	Upper Bound Kullback-Leibler Divergence for Hidden Markov Models with Application as Discrimination Measure for Speech Recognition. , 2006, , .		7
234	Analysis of Audio Clustering using Word Descriptions. , 2007, , .		7

#	ARTICLE	IF	CITATIONS
235	Data Driven Approach for Language Model Adaptation using Stepwise Relative Entropy Minimization. , 2007, , .		7
236	A semi-supervised learning approach to online audio background detection. , 2009, , .		7
237	Language model adaptation using WWW documents obtained by utterance-based queries. , 2010, , .		7
238	Using emotional noise to uncloud audio-visual emotion perceptual evaluation. , 2013, , .		7
239	Head motion synchrony and its correlation to affectivity in dyadic interactions. , 2013, , .		7
240	Power-spectral analysis of head motion signal for behavioral modeling in human interaction. , 2014, , .		7
241	Redundancy analysis of behavioral coding for couples therapy and improved estimation of behavior from noisy annotations. , 2015, , .		7
242	Opening big in box office? Trailer content can help. , 2016, , .		7
243	Directly data-derived articulatory gesture-like representations retain discriminatory information about phone categories. Computer Speech and Language, 2016, 36, 330-346.	4.3	7
244	Discovering Optimal Variable-length Time Series Motifs in Large-scale Wearable Recordings of Human Bio-behavioral Signals. , 2019, , .		7
245	An Empirical Study of Speech Processing in the Brain by Analyzing the Temporal Syllable Structure in Speech-input Induced EEG. , 2019, , .		7
246	Reinforcing Self-expressive Representation with Constraint Propagation for Face Clustering in Movies. , 2019, , .		7
247	Modeling Vocal Entrainment in Conversational Speech Using Deep Unsupervised Learning. IEEE Transactions on Affective Computing, 2022, 13, 1651-1663.	8.3	7
248	Meta-Learning for Robust Child-Adult Classification from Speech. , 2020, , .		7
249	Leveraging Linguistic Context in Dyadic Interactions to Improve Automatic Speech Recognition for Children. Computer Speech and Language, 2020, 63, 101101.	4.3	7
250	Deep multiple instance learning for foreground speech localization in ambient audio from wearable devices. Eurasip Journal on Audio, Speech, and Music Processing, 2021, 2021, 7.	2.1	7
251	End-to-end neural systems for automatic children speech recognition: An empirical study. Computer Speech and Language, 2022, 72, 101289.	4.3	7
252	Multi-Task Discriminative Training of Hybrid DNN-TVM Model for Speaker Verification with Noisy and Far-Field Speech. , 0, , .		7

#	ARTICLE	IF	CITATIONS
253	Advances in vocal tract imaging and analysis. , 2019, , 34-50.		7
254	Joint-processing of audio-visual signals in human perception of conflicting synthetic character emotions. , 2008, , .		6
255	Emotion classification from speech using evaluator reliability-weighted combination of ranked lists. , 2011, , .		6
256	A study on the effect of prosodic emphasis transfer on overall speech translation quality. , 2013, , .		6
257	Semi-supervised term-weighted value rescoring for keyword search. , 2014, , .		6
258	Graph-based approach for motion capture data representation and analysis. , 2014, , .		6
259	Multimodal Interaction Modeling of Child Forensic Interviewing. , 2018, , .		6
260	Improving the Prediction of Therapist Behaviors in Addiction Counseling by Exploiting Class Confusions. , 2019, , .		6
261	Breathing Rate Complexity Features for "In-the-Wild" Stress and Anxiety Measurement. , 2019, , .		6
262	Affect Estimation with Wearable Sensors. Journal of Healthcare Informatics Research, 2020, 4, 261-294.	7.6	6
263	Improving speaker diarization for naturalistic child-adult conversational interactions using contextual information. Journal of the Acoustical Society of America, 2020, 147, EL196-EL200.	1.1	6
264	Aliasing artifact reduction in spiral real-time MRI. Magnetic Resonance in Medicine, 2021, 86, 916-925.	3.0	6
265	Complexity of vocal tract shaping in glossectomy patients and typical speakers: A principal component analysis. Journal of the Acoustical Society of America, 2021, 149, 4437-4449.	1.1	6
266	Knowledge and Attitudes Toward an Artificial Intelligence-Based Fidelity Measurement in Community Cognitive Behavioral Therapy Supervision. Administration and Policy in Mental Health and Mental Health Services Research, 2021, , 1.	2.1	6
267	Engineering Innovation in Speech Science: Data and Technologies. Perspectives of the ASHA Special Interest Groups, 2019, 4, 411-420.	0.8	6
268	An investigation of vocal arousal dynamics in child-psychologist interactions using synchrony measures and a conversation-based model. , 0, , .		6
269	Combined Speaker Clustering and Role Recognition in Conversational Speech. , 0, , .		6
270	Factored translation models for enriching spoken language translation with prosody. , 0, , .		6

#	ARTICLE	IF	CITATIONS
271	Pronunciation verification of English letter-sounds in preliterate children. , 0, , .		6
272	The Second DIHARD Challenge: System Description for USC-SAIL Team. , 0, , .		6
273	Minimum Probability of Error Signal Representation. IEEE International Workshop on Machine Learning for Signal Processing, 2007, , .	0.0	5
274	The SAIL speaker diarization system for analysis of spontaneous meetings. , 2008, , .		5
275	Recognition for synthesis: Automatic parameter selection for resynthesis of emotional speech from neutral speech. Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, 2008, , .	1.8	5
276	An analysis of articulatory-acoustic data based on articulatory strokes. , 2009, , .		5
277	Automatic pronunciation verification of english letter-names for early literacy assessment of preliterate children. , 2009, , .		5
278	Accelerated 3D MRI of vocal tract shaping using compressed sensing and parallel imaging. , 2009, , .		5
279	Towards modeling user behavior in interactions mediated through an automated bidirectional speech translation system. Computer Speech and Language, 2010, 24, 232-256.	4.3	5
280	An audio-visual approach to learning salient behaviors in couples' problem solving discussions. , 2013, , .		5
281	On smoothing articulatory trajectories obtained from Gaussian mixture model based acoustic-to-articulatory inversion. Journal of the Acoustical Society of America, 2013, 134, EL258-EL264.	1.1	5
282	Gesture dynamics modeling for attitude analysis using graph based transform. , 2014, , .		5
283	Barista: A framework for concurrent speech processing by usc-sail. , 2014, 2014, 3306-3310.		5
284	Multiple Instance Learning for Behavioral Coding. IEEE Transactions on Affective Computing, 2017, 8, 81-94.	8.3	5
285	The Promise and the Challenge of Technology-Facilitated Methods for Assessing Behavioral and Cognitive Markers of Risk for Suicide among U.S. Army National Guard Personnel. International Journal of Environmental Research and Public Health, 2017, 14, 361.	2.6	5
286	Multimodal Representation of Advertisements Using Segment-level Autoencoders. , 2018, , .		5
287	Bluetooth Based Indoor Localization Using Triplet Embeddings. , 2019, , .		5
288	Stress and Anxiety Measurement "In-the-Wild" Using Quality-aware Multi-scale HRV Features. , 2019, 2019, 7056-7059.		5

#	ARTICLE	IF	CITATIONS
289	Modeling Behavioral Consistency in Large-Scale Wearable Recordings of Human Bio-Behavioral Signals. , 2020, , .		5
290	Modeling Behavior as Mutual Dependency between Physiological Signals and Indoor Location in Large-Scale Wearable Sensor Study. , 2020, , .		5
291	Vocal Tract Articulatory Contour Detection in Real-Time Magnetic Resonance Images Using Spatio-Temporal Context. , 2020, , .		5
292	Adversarial Defense for Deep Speaker Recognition Using Hybrid Adversarial Training. , 2021, , .		5
293	Movie Content Analysis, Indexing and Skimming Via Multimodal Information. , 2003, , 123-154.		5
294	Acoustic analysis and automatic recognition of spontaneous children's speech. , 0, , .		5
295	Multi-Scale Context Adaptation for Improving Child Automatic Speech Recognition in Child-Adult Spoken Interactions. , 0, , .		5
296	Identifying Therapist and Client Personae for Therapeutic Alliance Estimation. , 0, , .		5
297	Bilabial Substitution Patterns during Consonant Production in a Case of Congenital Aglossia. Journal of Communication Disorders Deaf Studies & Hearing Aids, 2017, 05, .	0.2	5
298	Towards an Unsupervised Entrainment Distance in Conversational Speech Using Deep Neural Networks. , 0, , .		5
299	Analyzing the Language of Therapist Empathy in Motivational Interview based Psychotherapy. , 2012, 2012, .		5
300	Identification of speakers in movie dialogs using audiovisual cues. , 2002, , .		4
301	ACOUSTIC-SYNTACTIC MAXIMUM ENTROPY MODEL FOR AUTOMATIC PROSODY LABELING. , 2006, , .		4
302	Multimodal Meeting Monitoring: Improvements on Speaker Tracking and Segmentation through a Modified Mixture Particle Filter. , 2007, , .		4
303	Bilingual audio-subtitle extraction using automatic segmentation of movie audio. , 2011, , .		4
304	Supervised acoustic topic model with a consequent classifier for unstructured audio classification. , 2012, , .		4
305	On signal representations within the Bayes decision framework. Pattern Recognition, 2012, 45, 1853-1865.	8.1	4
306	Toward body language generation in dyadic interaction settings from interlocutor multimodal cues. , 2013, , .		4

#	ARTICLE	IF	CITATIONS
307	Spatial and temporal alignment of multimodal human speech production data: Real time imaging, flesh point tracking and audio. , 2013, , .		4
308	Modeling mutual influence of multimodal behavior in affective dyadic interactions. , 2015, , .		4
309	Weighted geodesic flow kernel for interpersonal mutual influence modeling and emotion recognition in dyadic interactions. , 2017, , .		4
310	Discovering Latent Psychological Structures from Self-Report Assessments of Hospital Workers. , 2018, , .		4
311	Participatory methods to support team science development for predictive analytics in health. Journal of Clinical and Translational Science, 2018, 2, 178-182.	0.6	4
312	Total Variability Layer in Deep Neural Network Embeddings for Speaker Verification. IEEE Signal Processing Letters, 2019, 26, 893-897.	3.6	4
313	Predicting Human-Reported Enjoyment Responses in Happy and Sad Music. , 2019, , .		4
314	L2 Acquisition and Production of the English Rhotic Pharyngeal Gesture. , 0, , .		4
315	Extracting Situation Frames from Non-English Speech: Evaluation Framework and Pilot Results. , 0, , .		4
316	Computational Modeling of Conversational Humor in Psychotherapy. , 2018, 2018, 2344-2348.		4
317	Multiview Shared Subspace Learning Across Speakers and Speech Commands. , 0, , .		4
318	User-Based Collaborative Filtering Mobile Health System. , 2020, 4, 1-17.		4
319	An automated quality evaluation framework of psychotherapy conversations with local quality estimates. Computer Speech and Language, 2022, 75, 101380.	4.3	4
320	Discriminating Two Types of Noise Sources using Cortical Representation and Dimension Reduction Technique. , 2007, , .		3
321	Experiments in Automatic Genre Classification of Full-length Music Tracks using Audio Activity Rate. , 2007, , .		3
322	Histogram-based estimation for the divergence revisited. , 2009, , .		3
323	Robust representations for out-of-domain emotions using Emotion Profiles. , 2010, , .		3
324	Bark Frequency Transform Using an Arbitrary Order Allpass Filter. IEEE Signal Processing Letters, 2010, 17, 543-546.	3.6	3

#	ARTICLE	IF	CITATIONS
325	Automatically assessing the ABCs. ACM Transactions on Speech and Language Processing, 2011, 7, 1-17.	0.9	3
326	Improvements in predicting children's overall reading ability by modeling variability in evaluators' subjective judgments. , 2012, , .		3
327	Creating ensemble of diverse maximum entropy models. , 2012, , .		3
328	A robust frontend for ASR: Combining denoising, noise masking and feature normalization. , 2013, , .		3
329	Energy-constrained minimum variance response filter for robust vowel spectral estimation. , 2014, , .		3
330	Fusion of diverse denoising systems for robust automatic speech recognition. , 2014, , .		3
331	The ELISA Situation Frame extraction for low resource languages pipeline for LoReHLT™2016. Machine Translation, 2018, 32, 127-142.	1.3	3
332	Generating labels for regression of subjective constructs using triplet embeddings. Pattern Recognition Letters, 2019, 128, 385-392.	4.2	3
333	Prediction of Psychological Flexibility with multi-scale Heart Rate Variability and Breathing Features in an "in-the-wild" Setting. , 2019, , .		3
334	Subspace techniques for task-independent EEG person identification. , 2019, 2019, 4545-4548.		3
335	Variability in individual constriction contributions to third formant values in American English /É¹/. Journal of the Acoustical Society of America, 2020, 147, 3905-3916.	1.1	3
336	Robust Character Labeling in Movie Videos: Data Resources and Self-Supervised Feature Adaptation. IEEE Transactions on Multimedia, 2022, 24, 3355-3368.	7.2	3
337	Multi-Scale Speaker Diarization with Neural Affinity Score Fusion. , 2021, , .		3
338	Temporal Dynamics of Workplace Acoustic Scenes: Egocentric Analysis and Prediction. IEEE/ACM Transactions on Audio Speech and Language Processing, 2021, 29, 756-769.	5.8	3
339	EmotiWord: Affective Lexicon Creation with Application to Interaction and Multimedia Data. Lecture Notes in Computer Science, 2012, , 30-41.	1.3	3
340	Improved Depiction of Tissue Boundaries in Vocal Tract Real-Time MRI Using Automatic Off-Resonance Correction. , 0, , .		3
341	Mitigating the Bias of Heterogeneous Human Behavior in Affective Computing. , 2021, , .		3
342	Intra-topic latency as an automated behavioral marker of treatment response in autism spectrum disorder. Scientific Reports, 2022, 12, 3255.	3.3	3

#	ARTICLE	IF	CITATIONS
343	Variation in compensatory strategies as a function of target constriction degree in post-glossectomy speech. <i>JASA Express Letters</i> , 2022, 2, 045205.	1.1	3
344	An attribute-based approach to audio description applied to segmenting vocal sections in popular music songs. , 2006, , .		2
345	Selection of Emotionally Salient Audio-Visual Features for Modeling Human Evaluations of Synthetic Character Emotion Displays. , 2008, , .		2
346	Overlapped speech detection using long-term spectro-temporal similarity in stereo recording. , 2011, , .		2
347	Features for comparing tune similarity of songs across different languages. , 2012, , .		2
348	Classification of emotional content of sighs in dyadic human interactions. , 2012, , .		2
349	Enriching machine-mediated speech-to-speech translation using contextual information. <i>Computer Speech and Language</i> , 2013, 27, 492-508.	4.3	2
350	Integration and Automation of Data Preparation and Data Mining. , 2014, , .		2
351	Learning multiple concepts with incremental diverse density. , 2014, , .		2
352	Affect prediction in music using boosted ensemble of filters. , 2015, , .		2
353	Lightly-supervised utterance-level emotion identification using latent topic modeling of multimodal words. , 2016, , .		2
354	Improving Semi-Supervised Classification for Low-Resource Speech Interaction Applications. , 2018, , .		2
355	Efficient estimation and model generalization for the totalvariability model. <i>Computer Speech and Language</i> , 2019, 53, 43-64.	4.3	2
356	Learning Shared Vector Representations of Lyrics and Chords in Music. , 2019, , .		2
357	Cross-Modal Coordination of Face-Directed Gaze and Emotional Speech Production in School-Aged Children and Adolescents with ASD. <i>Scientific Reports</i> , 2019, 9, 18301.	3.3	2
358	Improved real-time tagged MRI using REALTAG. <i>Magnetic Resonance in Medicine</i> , 2020, 84, 838-846.	3.0	2
359	Identifying Truthful Language in Child Interviews. , 2020, , .		2
360	Developing Neural Representations for Robust Child-Adult Diarization. , 2021, , .		2

#	ARTICLE	IF	CITATIONS
361	Causal indicators for assessing the truthfulness of child speech in forensic interviews. <i>Computer Speech and Language</i> , 2022, 71, 101263.	4.3	2
362	Automated Empathy Detection for Oncology Encounters. , 2020, , .		2
363	Novel affective features for multiscale prediction of emotion in music. , 2016, , .		2
364	Perceptual Lateralization of Coda Rhotic Production in Puerto Rican Spanish. , 0, , .		2
365	A Distribution Free Formulation of the Total Variability Model. , 0, , .		2
366	A Knowledge Driven Structural Segmentation Approach for Play-Talk Classification During Autism Assessment. , 0, , .		2
367	Stochastic Shake-Shake Regularization for Affective Learning from Speech. , 0, , .		2
368	Privacy and Utility Preserving Data Transformation for Speech Emotion Recognition. , 2021, , .		2
369	Automatic Analysis of Asymmetry in Facial Paralysis Patients Using Landmark-Based Measures. <i>Facial Plastic Surgery and Aesthetic Medicine</i> , 2022, 24, 491-493.	0.9	2
370	A divide-and-conquer approach to Latent Perceptual Indexing of audio for large Web 2.0 applications. , 2009, , .		1
371	A robust harmony structure modeling scheme for classical music opus identification. , 2009, , .		1
372	Lattice-based lexical cues for word fragment detection in conversational speech. , 2009, , .		1
373	Using naïve text queries for robust audio information retrieval. , 2010, , .		1
374	High-quality bilingual subtitle document alignments with application to spontaneous speech translation. <i>Computer Speech and Language</i> , 2013, 27, 572-591.	4.3	1
375	Combining window predictions efficiently - A new imputation approach for noise robust automatic speech recognition. , 2013, , .		1
376	Hull detection based on largest empty sector angle with application to analysis of realtime MR images. , 2014, , .		1
377	Comparison of feature-level and kernel-level data fusion methods in multi-sensory fall detection. , 2016, , .		1
378	Speech and language processing for mental health research and care. , 2016, , .		1

#	ARTICLE	IF	CITATIONS
379	A knowledge transfer and boosting approach to the prediction of affect in movies. , 2017, , .		1
380	Role Annotated Speech Recognition for Conversational Interactions. , 2018, , .		1
381	Role Specific Lattice Rescoring for Speaker Role Recognition from Speech Recognition Outputs. , 2019, , .		1
382	Trapezoidal Segmented Regression: A Novel Continuous-scale Real-time Annotation Approximation Algorithm. , 2019, , .		1
383	A system for the 2019 Sentiment, Emotion and Cognitive State Task of DARPA's LORELEI project. , 2019, , .		1
384	Joint Multi-Dimensional Model for Global and Time-Series Annotations. IEEE Transactions on Affective Computing, 2022, 13, 473-484.	8.3	1
385	Learning Domain Invariant Representations for Child-Adult Classification from Speech. , 2020, , .		1
386	Bringing in the Outliers: A Sparse Subspace Clustering Approach to Learn a Dictionary of Mouse Ultrasonic Vocalizations. , 2020, , .		1
387	The Role of Annotation Fusion Methods in the Study of Human-Reported Emotion Experience During Music Listening. , 2020, , .		1
388	Conversational correlates of rapid social judgments of children and adolescents with and without ASD. Clinical Linguistics and Phonetics, 2021, 35, 172-184.	0.9	1
389	Generalized Multiview Shared Subspace Learning Using View Bootstrapping. IEEE Transactions on Signal Processing, 2021, 69, 4774-4786.	5.3	1
390	Dark tone quality and vocal tract shaping in soprano song production: Insights from real-time MRI. JASA Express Letters, 2021, 1, 075202.	1.1	1
391	How an aglossic speaker produces an alveolar-like percept without a functional tongue tip. Journal of the Acoustical Society of America, 2020, 147, EL460-EL464.	1.1	1
392	Exploring the Relationship between Conic Affinity of NMF Dictionaries and Speech Enhancement Metrics. , 0, , .		1
393	Denoising and Raw-waveform Networks for Weakly-Supervised Gender Identification on Noisy Speech. , 0, , .		1
394	Confusion2Vec 2.0: Enriching ambiguous spoken language representations with subwords. PLoS ONE, 2022, 17, e0264488.	2.5	1
395	Representation of professions in entertainment media: Insights into frequency and sentiment trends through computational text analysis. PLoS ONE, 2022, 17, e0267812.	2.5	1
396	Optimal Wavelet Packets Decomposition Based on a Rate-Distortion Optimality Criterion. , 2007, , .		0

#	ARTICLE	IF	CITATIONS
397	Analyzing the Multimodal Behaviors of Users of a Speech-to-Speech Translation Device by using Concept Matching Scores. , 2007, , .		0
398	Incorporating discourse context in spoken language translation through dialog acts. , 2008, , .		0
399	A near-optimal (minimax) tree-structured partition for mutual information estimation. , 2010, , .		0
400	On data-driven histogram-based estimation for mutual information. , 2010, , .		0
401	Auditory-like filterbank: An optimal speech processor for efficient human speech communication. Sadhana - Academy Proceedings in Engineering Sciences, 2011, 36, 699-712.	1.3	0
402	Comparison of feature-level and kernel-level data fusion methods in multi-sensory fall detection. , 2016, , .		0
403	On Role and Location of Normalization before Model-based Data Augmentation in Residual Blocks for Classification Tasks. , 2019, , .		0
404	Trapezoidal Segment Sequencing: A Novel Approach for Fusion of Human-Produced Continuous Annotations. , 2020, , .		0
405	A Computational Tool to Study Vocal Participation of Women in UN-ITU Meetings. , 2021, , .		0
406	Test-Retest Repeatability of Articulatory Strategies Using Real-Time Magnetic Resonance Imaging. , 0, , .		0
407	Exploiting Intra-Annotator Rating Consistency Through Copelandâ€™s Method for Estimation of Ground Truth Labels in Couplesâ€™ Therapy. , 0, , .		0
408	Itâ€™s not what you said, itâ€™s how you said it: An analysis of therapist vocal features during psychotherapy. Counselling and Psychotherapy Research, 2023, 23, 258-269.	3.2	0