Carlos Busso

List of Publications by Year
in descending order

156 papers	7,589 citations	201575 27 h-index	168321 53 g-index
156 all docs	156 docs citations	156 times ranked	3434 citing authors

#	Article	IF	Citations
1	Chunk-Level Speech Emotion Recognition: A General Framework of Sequence-to-One Dynamic Temporal Modeling. IEEE Transactions on Affective Computing, 2023, 14, 1215-1227.	5.7	13
2	Quantifying Emotional Similarity in Speech. IEEE Transactions on Affective Computing, 2023, 14, 1376-1390.	5.7	5
3	Estimation of Driver's Gaze Region From Head Position and Orientation Using Probabilistic Confidence Regions. IEEE Transactions on Intelligent Vehicles, 2023, 8, 59-72.	9.4	18
4	Temporal Head Pose Estimation From Point Cloud in Naturalistic Driving Conditions. IEEE Transactions on Intelligent Transportation Systems, 2022, 23, 8063-8076.	4.7	8
5	The Multimodal Driver Monitoring Database: A Naturalistic Corpus to Study Driver Attention. IEEE Transactions on Intelligent Transportation Systems, 2022, 23, 10736-10752.	4.7	8
6	Not All Features are Equal: Selection of Robust Features for Speech Emotion Recognition in Noisy Environments. , 2022, , .		6
7	Incorporating Gaze Behavior Using Joint Embedding With Scene Context for Driver Takeover Detection. , 2022, , .		2
8	Exploiting Annotators' Typed Description of Emotion Perception to Maximize Utilization of Ratings for Speech Emotion Recognition. , 2022, , .		7
9	AuxFormer: Robust Approach to Audiovisual Emotion Recognition. , 2022, , .		6
10	Face detection and grimace scale prediction of white furred mice. Machine Learning With Applications, 2022, 8, 100312.	3.0	5
11	Unsupervised Personalization of an Emotion Recognition System: The Unique Properties of the Externalization of Valence in Speech. IEEE Transactions on Affective Computing, 2022, 13, 1959-1972.	5.7	5
12	Speech-Driven Expressive Talking Lips with Conditional Sequential Generative Adversarial Networks. IEEE Transactions on Affective Computing, 2021, 12, 1031-1044.	5.7	23
13	Over-Sampling Emotional Speech Data Based on Subjective Evaluations Provided by Multiple Individuals. IEEE Transactions on Affective Computing, 2021, 12, 870-882.	5.7	6
14	End-to-End Audiovisual Speech Recognition System With Multitask Learning. IEEE Transactions on Multimedia, 2021, 23, 1-11.	5.2	54
15	The Ordinal Nature of Emotions: An Emerging Approach. IEEE Transactions on Affective Computing, 2021, 12, 16-35.	5.7	55
16	Predicting Emotionally Salient Regions Using Qualitative Agreement of Deep Neural Network Regressors. IEEE Transactions on Affective Computing, 2021, 12, 402-416.	5.7	8
17	Smartphone sensing of social interactions in people with and without schizophrenia. Journal of Psychiatric Research, 2021, 137, 613-620.	1.5	39
18	Guided Generative Adversarial Neural Network for Representation Learning and Audio Generation Using Fewer Labelled Audio Data. IEEE/ACM Transactions on Audio Speech and Language Processing, 2021, 29, 2575-2590.	4.0	5

#	Article	IF	Citations
19	Deepemocluster: a Semi-Supervised Framework for Latent Cluster Representation of Speech Emotions. , 2021, , .		6
20	Machine-Learning Assisted Discrimination of Precancerous and Cancerous from Healthy Oral Tissue Based on Multispectral Autofluorescence Lifetime Imaging Endoscopy. Cancers, 2021, 13, 4751.	1.7	19
21	Multimodal Behavior Modeling for Socially Interactive Agents. , 2021, , 259-310.		10
22	Deep Representation Learning for Affective Speech Signal Analysis and Processing: Preventing unwanted signal disparities. IEEE Signal Processing Magazine, 2021, 38, 22-38.	4.6	16
23	Generative Approach Using Soft-Labels to Learn Uncertainty in Predicting Emotional Attributes. , 2021, , .		6
24	Semi-Supervised Speech Emotion Recognition With Ladder Networks. IEEE/ACM Transactions on Audio Speech and Language Processing, 2020, 28, 2697-2709.	4.0	66
25	Style Extractor For Facial Expression Recognition in the Presence of Speech. , 2020, , .		2
26	8 Head Pose as an Indicator of Drivers' Visual Attention. , 2020, , 113-132.		3
27	Modeling Uncertainty in Predicting Emotional Attributes from Spontaneous Speech. , 2020, , .		9
28	Use of Triplet-Loss Function to Improve Driving Anomaly Detection Using Conditional Generative Adversarial Network. , 2020, , .		8
29	MSP-Face Corpus. , 2020, , .		7
30	Robust Driver Head Pose Estimation in Naturalistic Conditions from Point-Cloud Data. , 2020, , .		10
31	Dynamic versus Static Facial Expressions in the Presence of Speech. , 2020, , .		1
32	End-to-end audiovisual speech activity detection with bimodal recurrent neural models. Speech Communication, 2019, 113, 25-35.	1.6	18
33	Driving Anomaly Detection with Conditional Generative Adversarial Network using Physiological and CAN-Bus Data. , 2019, , .		18
34	Discriminative Features for Texture Retrieval Using Wavelet Packets. IEEE Access, 2019, 7, 148882-148896.	2.6	3
35	Retrieving Speech Samples with Similar Emotional Content Using a Triplet Loss Function. , 2019, , .		9
36	Speech-driven animation with meaningful behaviors. Speech Communication, 2019, 110, 90-100.	1.6	27

#	ARTICLE	IF	CITATIONS
37	Estimation of Gaze Region Using Two Dimensional Probabilistic Maps Constructed Using Convolutional Neural Networks. , 2019, , .		6
38	Lexical Dependent Emotion Detection Using Synthetic Speech Reference. IEEE Access, 2019, 7, 22071-22085.	2.6	10
39	Curriculum Learning for Speech Emotion Recognition From Crowdsourced Labels. IEEE/ACM Transactions on Audio Speech and Language Processing, 2019, 27, 815-826.	4.0	57
40	Exploring the Intersection Between Speaker Verification and Emotion Recognition. , 2019, , .		4
41	Active Learning for Speech Emotion Recognition Using Deep Neural Network. , 2019, , .		20
42	Analysis of the Relationship Between Physiological Signals and Vehicle Maneuvers During a Naturalistic Driving Study. , 2019, , .		8
43	Building Naturalistic Emotionally Balanced Speech Corpus by Retrieving Emotional Speech from Existing Podcast Recordings. IEEE Transactions on Affective Computing, 2019, 10, 471-483.	5.7	119
44	Calibration free, user-independent gaze estimation with tensorÂanalysis. Image and Vision Computing, 2018, 74, 10-20.	2.7	7
45	Gating Neural Network for Large Vocabulary Audiovisual Speech Recognition. IEEE/ACM Transactions on Audio Speech and Language Processing, 2018, 26, 1290-1302.	4.0	42
46	Aligning Audiovisual Features for Audiovisual Speech Recognition. , 2018, , .		16
47	Novel Realizations of Speech-Driven Head Movements with Generative Adversarial Networks. , 2018, , .		34
48	Probabilistic Estimation of the Gaze Region of the Driver using Dense Classification. , 2018, , .		22
49	FI-CAP: Robust Framework to Benchmark Head Pose Estimation in Challenging Environments. , 2018, , .		5
50	Study of Dense Network Approaches for Speech Emotion Recognition. , 2018, , .		27
51	Expressive Speech-Driven Lip Movements with Multitask Learning. , 2018, , .		9
52	Domain Adversarial for Acoustic Emotion Recognition. IEEE/ACM Transactions on Audio Speech and Language Processing, 2018, 26, 2423-2435.	4.0	110
53	Head Motion Generation. , 2018, , 2177-2200.		2
54	MSP-IMPROV: An Acted Corpus of Dyadic Interactions to Study Emotion Perception. IEEE Transactions on Affective Computing, 2017, 8, 67-80.	5.7	192

#	Article	IF	Citations
55	The Cost of Dichotomizing Continuous Labels for Binary Classification Problems: Deriving a Bayesian-Optimal Classifier. IEEE Transactions on Affective Computing, 2017, 8, 119-130.	5.7	13
56	4. Driver mirror-checking action detection. , 2017, , 55-76.		1
57	Meaningful head movements driven by emotional synthetic speech. Speech Communication, 2017, 95, 87-99.	1.6	17
58	Driver Modeling for Detection and Assessment of Driver Distraction: Examples from the UTDrive Test Bed. IEEE Signal Processing Magazine, 2017, 34, 130-142.	4.6	47
59	A study of speaker verification performance with expressive speech. , 2017, , .		10
60	Ensemble feature selection for domain adaptation in speech emotion recognition. , 2017, , .		11
61	Predicting speaker recognition reliability by considering emotional content. , 2017, , .		8
62	The ordinal nature of emotions. , 2017, , .		75
63	Challenges in head pose estimation of drivers in naturalistic recordings using existing tools. , 2017, , .		15
64	Probabilistic estimation of the driver's gaze from head orientation and position. , 2017, , .		7
65	Formulating emotion perception as a probabilistic model with application to categorical emotion classification. , 2017, , .		18
66	Assessment and classification of singing quality based on audio-visual features. , 2017, , .		0
67	Incremental adaptation using active learning for acoustic emotion recognition. , 2017, , .		15
68	Joint Learning of Speech-Driven Facial Motion with Bidirectional Long-Short Term Memory. Lecture Notes in Computer Science, 2017, , 389-402.	1.0	8
69	Ranking emotional attributes with deep neural networks. , 2017, , .		19
70	Tradeoff between quality and quantity of emotional annotations to characterize expressive behaviors. , 2016, , .		10
71	Analyzing the relationship between head pose and gaze to model driver visual attention. , 2016, , .		24
72	Using Agreement on Direction of Change to Build Rank-Based Emotion Classifiers. IEEE/ACM Transactions on Audio Speech and Language Processing, 2016, 24, 2108-2121.	4.0	21

#	Article	IF	Citations
73	Automatic composition of broadcast news summaries using rank classifiers trained with acoustic and lexical features. , 2016, , .		2
74	Practical considerations on the use of preference learning for ranking emotional speech. , 2016, , .		25
75	A multimodal analysis of synchrony during dyadic interaction using a metric based on sequential pattern mining. , 2016, , .		4
76	Detecting Drivers' Mirror-Checking Actions and Its Application to Maneuver and Secondary Task Recognition. IEEE Transactions on Intelligent Transportation Systems, 2016, 17, 980-992.	4.7	64
77	Increasing the Reliability of Crowdsourcing Evaluations Using Online Quality Assessment. IEEE Transactions on Affective Computing, 2016, 7, 374-388.	5.7	75
78	The Geneva Minimalistic Acoustic Parameter Set (GeMAPS) for Voice Research and Affective Computing. IEEE Transactions on Affective Computing, 2016, 7, 190-202.	5.7	926
79	Facial Expression Recognition in the Presence of Speech Using Blind Lexical Compensation. IEEE Transactions on Affective Computing, 2016, 7, 346-359.	5.7	21
80	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.	1.8	41
81	Head Motion Generation. , 2016, , 1-25.		1
82	Retrieving Target Gestures Toward Speech Driven Animation with Meaningful Behaviors. , 2015, , .		12
83	Speech in Affective Computing. , 2015, , .		17
84	Correcting Time-Continuous Emotional Labels by Modeling the Reaction Lag of Evaluators. IEEE Transactions on Affective Computing, 2015, 6, 97-108.	5.7	77
85	Adjacent Vehicle Collision Warning System using Image Sensor and Inertial Measurement Unit. , 2015, , .		2
86	MSP-AVATAR corpus: Motion capture recordings to study the role of discourse functions in the design of intelligent virtual agents. , 2015, , .		12
87	Predicting Perceived Visual and Cognitive Distractions of Drivers With Multimodal Features. IEEE Transactions on Intelligent Transportation Systems, 2015, 16, 51-65.	4.7	71
88	UMEME: University of Michigan Emotional McGurk Effect Data Set. IEEE Transactions on Affective Computing, 2015, 6, 395-409.	5.7	19
89	Challenges in Concussion Detection Using Vocal Acoustic Biomarkers. IEEE Access, 2015, 3, 1143-1160.	2.6	13
90	Emotion recognition using synthetic speech as neutral reference. , 2015, , .		12

#	Article	IF	Citations
91	Supervised domain adaptation for emotion recognition from speech. , 2015, , .		45
92	Evaluation of syllable rate estimation in expressive speech and its contribution to emotion recognition. , 2014, , .		6
93	Speech-Driven Animation Constrained by Appropriate Discourse Functions. , 2014, , .		14
94	User Independent Gaze Estimation by Exploiting Similarity Measures in the Eye Pair Appearance Eigenspace. , 2014, , .		3
95	Compensating for speaker or lexical variabilities in speech for emotion recognition. Speech Communication, 2014, 57, 1-12.	1.6	43
96	Shape-based modeling of the fundamental frequency contour for emotion detection in speech. Computer Speech and Language, 2014, 28, 278-294.	2.9	45
97	Using Perceptual Evaluation to Quantify Cognitive and Visual Driver Distractions. , 2014, , 183-207.		4
98	Exploring Cross-Modality Affective Reactions for Audiovisual Emotion Recognition. IEEE Transactions on Affective Computing, 2013, 4, 183-196.	5.7	66
99	Analysis of facial features of drivers under cognitive and visual distractions. , 2013, , .		31
100	Modeling of Driver Behavior in Real World Scenarios Using Multiple Noninvasive Sensors. IEEE Transactions on Multimedia, 2013, 15, 1213-1225.	5.2	68
101	Feature and model level compensation of lexical content for facial emotion recognition. , 2013, , .		17
102	Analysis and Compensation of the Reaction Lag of Evaluators in Continuous Emotional Annotations. , 2013, , .		37
103	Audiovisual corpus to analyze whisper speech. , 2013, , .		13
104	Iterative Feature Normalization Scheme for Automatic Emotion Detection from Speech. IEEE Transactions on Affective Computing, 2013, 4, 386-397.	5.7	46
105	Evaluating the robustness of an appearance-based gaze estimation method for multimodal interfaces. , 2013, , .		7
106	Toward Effective Automatic Recognition Systems of Emotion in Speech. , 2013, , 110-127.		58
107	Generating Human-Like Behaviors Using Joint, Speech-Driven Models for Conversational Agents. IEEE Transactions on Audio Speech and Language Processing, 2012, 20, 2329-2340.	3.8	47
108	Factorizing speaker, lexical and emotional variabilities observed in facial expressions. , 2012, , .		11

#	Article	IF	Citations
109	A personalized emotion recognition system using an unsupervised feature adaptation scheme. , 2012, , .		24
110	Indoor robotic terrain classification via angular velocity based hierarchical classifier selection. , 2012, , .		15
111	Advances in Multimodal Tracking of Driver Distraction. , 2012, , 253-270.		11
112	Emotion recognition using a hierarchical binary decision tree approach. Speech Communication, 2011, 53, 1162-1171.	1.6	274
113	Analysis of driver behaviors during common tasks using frontal video camera and CAN-Bus information. , 2011, , .		19
114	Iterative feature normalization for emotional speech detection. , 2011, , .		40
115	Visual emotion recognition using compact facial representations and viseme information. , 2010, , .		40
116	Analysis of Emotionally Salient Aspects of Fundamental Frequency for Emotion Detection. IEEE Transactions on Audio Speech and Language Processing, 2009, 17, 582-596.	3.8	211
117	Interpreting ambiguous emotional expressions. , 2009, , .		89
118	IEMOCAP: interactive emotional dyadic motion capture database. Language Resources and Evaluation, 2008, 42, 335-359.	1.8	1,756
119	Learning Expressive Human-Like Head Motion Sequences from Speech. , 2008, , 113-131.		11
120	Real-Time Monitoring of Participants' Interaction in a Meeting using Audio-Visual Sensors. , 2007, , .		18
121	Multimodal Meeting Monitoring: Improvements on Speaker Tracking and Segmentation through a Modified Mixture Particle Filter. , 2007, , .		4
122	Joint Analysis of the Emotional Fingerprint in the Face and Speech: A single subject study. , 2007, , .		14
123	Rigid Head Motion in Expressive Speech Animation: Analysis and Synthesis. IEEE Transactions on Audio Speech and Language Processing, 2007, 15, 1075-1086.	3.8	147
124	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.	3.8	101
125	Natural head motion synthesis driven by acoustic prosodic features. Computer Animation and Virtual Worlds, 2005, 16, 283-290.	0.7	71
126	Audio-based head motion synthesis for Avatar-based telepresence systems. , 2004, , .		25

#	Article	IF	Citations
127	Analysis of emotion recognition using facial expressions, speech and multimodal information. , 2004, , .		555
128	Emotion recognition based on phoneme classes. , 0, , .		111
129	An unsupervised visual-only voice activity detection approach using temporal orofacial features. , 0, , .		6
130	Improving Boundary Estimation in Audiovisual Speech Activity Detection Using Bayesian Information Criterion. , 0, , .		6
131	Head Motion Generation with Synthetic Speech: A Data Driven Approach. , 0, , .		8
132	A Portable Automatic PA-TA-KA Syllable Detection System to Derive Biomarkers for Neurological Disorders. , 0, , .		7
133	A Stepwise Analysis of Aggregated Crowdsourced Labels Describing Multimodal Emotional Behaviors. , 0, , .		7
134	Jointly Predicting Arousal, Valence and Dominance with Multi-Task Learning. , 0, , .		79
135	Bimodal Recurrent Neural Network for Audiovisual Voice Activity Detection. , 0, , .		12
136	Ladder Networks for Emotion Recognition: Using Unsupervised Auxiliary Tasks to Improve Predictions of Emotional Attributes. , 0, , .		32
137	Predicting Categorical Emotions by Jointly Learning Primary and Secondary Emotions through Multitask Learning. , 0, , .		25
138	Role of Regularization in the Prediction of Valence from Speech. , 0, , .		14
139	An acoustic study of emotions expressed in speech. , 0, , .		84
140	Using neutral speech models for emotional speech analysis. , 0, , .		43
141	The expression and perception of emotions: comparing assessments of self versus others. , 0, , .		19
142	Scripted dialogs versus improvisation: lessons learned about emotional elicitation techniques from the IEMOCAP database. , 0, , .		17
143	Modeling mutual influence of interlocutor emotion states in dyadic spoken interactions. , 0, , .		38
144	Emotion recognition using a hierarchical binary decision tree approach. , 0, , .		37

#	Article	IF	Citations
145	Detecting sleepiness by fusing classifiers trained with novel acoustic features. , 0, , .		9
146	Unveiling the acoustic properties that describe the valence dimension. , 0, , .		20
147	Energy and F0 contour modeling with functional data analysis for emotional speech detection. , 0, , .		13
148	Lipreading approach for isolated digits recognition under whisper and neutral speech. , 0, , .		12
149	Defining Emotionally Salient Regions Using Qualitative Agreement Method. , 0, , .		13
150	Retrieving Categorical Emotions Using a Probabilistic Framework to Define Preference Learning Samples. , 0, , .		22
151	Preference-Learning with Qualitative Agreement for Sentence Level Emotional Annotations. , 0, , .		13
152	Audiovisual Speech Activity Detection with Advanced Long Short-Term Memory. , 0, , .		8
153	Speech Emotion Recognition with a Reject Option. , 0, , .		7
154	An Efficient Temporal Modeling Approach for Speech Emotion Recognition by Mapping Varied Duration Sentences into Fixed Number of Chunks. , 0, , .		12
155	Ensemble of Students Taught by Probabilistic Teachers to Improve Speech Emotion Recognition. , 0, , .		4
156	The MSP-Conversation Corpus. , 0, , .		11