

Carlos Busso

email: busso@utdallas.edu

website: www.utdallas.edu/~busso

phone: (972) 883-4351

RESEARCH INTERESTS	His research interest is in human-centered multimodal machine intelligence and applications. His current research includes the broad areas of speech processing, affective computing, multimodal human-machine interfaces, in-vehicle active safety systems, and machine learning methods for multimodal processing. His work has direct implications in many practical domains, including national security, health care, entertainment, transportation systems, and education.	
EDUCATION	University of Southern California , Los Angeles, CA	2003-2008
	Ph.D. in Electrical Engineering	
	Advisor: Shrikanth Narayanan <i>Niki & C. L. Max Nikias Chair in Engineering</i>	
	<i>Thesis:</i> “Multimodal Analysis of Expressive Human Communication: Speech and gesture interplay”	
	University of Chile , Santiago, Chile	2001-2003
	Master of Science in Electrical Engineering	
	Advisor: Néstor Becerra Yoma	
	<i>Thesis:</i> “Speech transmission over IP with a protocol based on the LMS algorithm” (Spanish)	
	University of Chile , Santiago, Chile	1996-2003
	Electrical Engineer	
	University of Chile , Santiago, Chile	1996-2000
	Bachelor in Electrical Engineering	
PROFESSIONAL EXPERIENCE	University of Texas at Dallas , Richardson, TX, USA	
	Professor	2020-present
	Associate Professor	2015-2020
	Assistant Professor	2009-2015
	Director of the Multimodal Signal Processing (MSP) lab [msp.utdallas.edu]	
	University of Southern California , Los Angeles, CA, USA	
	Postdoctoral Research Associate	2008-2009
	Mentor: Shrikanth Narayanan	
	Analysis of expressive human communication	
	Motorola , Austin, TX, USA	
	Summer Intern	Summer 2001
	Mentor: Kate Stewart	
	Optimized and modified C code (project Media Stream Resource Module)	
	Telefónica Móvil , Santiago, Chile	
	Summer Intern	Summer 2000
	Mentor: Andres Guerra	
	Analyzed and tested quality of service in wireless communication network	
	Adexus S.A. , Santiago, Chile	
	Summer Intern	Summer 1999
	Mentor: Aurelio Luco	
	Reported quality of service in optical communication networks	
HONORS, AWARDS AND FELLOWSHIPS	IEEE Fellow	2023
	ACM ICMI Community Service Award	2023
	SIPI Distinguished Alumni Award in the Mid-Career/Academia category	2023
	Federal Research Innovation and Expenditures Dynamite (FRIEND) at UTD	2023
	2021 Best Paper Award from IEEE Transactions on Affective Computing	2022

Best of IEEE Transactions on Affective Computing Paper Collection award	2021
Best Paper Award – AAAC Affective Computing and Intelligent Interaction (ACII)	2017
Recognition of Service Award from ACM	2016
Outstanding Reviewer for Visual Communications and Image Processing (VCIP)	2016
Jonsson School Senior (Assoc. Professor) Research Award	2016
IEEE Access Best Multimedia Contest Winner 2015	2015
NSF CAREER award	2015
Third prize IEEE ITSS Best Dissertation Award 2015 (student: Nanxiang Li)	2015
ICMI Ten-Year Technical Impact Award – International conference on multimodal interaction (ICMI)	2014
Quality Reviewer – IEEE International Conference on Multimedia and Expo (ICME)	2013
Recognition for Excellence in Reviewing – IEEE Conference on Automatic Face and Gesture Recognition (FG)	2013
Best Paper Award – IEEE International Conference on Multimedia and Expo (ICME)	2011
Quality Reviewer – IEEE International Conference on Multimedia and Expo (ICME)	2011
First place, Classifier Sub-Challenge event at the Interspeech 2009 emotion challenge	2009
Fellowship in Digital Scholarship, University of Southern California, Los Angeles, CA	2007-2008
Provost’s Fellowship, University of Southern California, Los Angeles, CA	2003-2005
Selected the best Electrical Engineer graduated in Chile by the Chilean School of Engineering	2003

PROFESSIONAL MEMBERSHIPS

Fellow of the Institute of Electrical and Electronics Engineers (IEEE)
IEEE Signal Processing Society
IEEE Computer Society
IEEE Intelligent Transportation Systems Society
Senior Member of the Association for Computing Machinery (ACM)
Member of the International Speech Communication Association (ISCA)
Member of the Association for the Advancement of Affective Computing (AAAC)

MOST RELEVANT PROFESSIONAL ACTIVITIES

- ★ Associated Editor for IEEE Transactions on Affective Computing (Mar.2020-present)
- ★ Senior editor of IEEE Signal Processing Letters (Aug. 2018-Jul. 2022)
- ★ Associated Editor for IEEE/ACM Transaction on Audio, Speech, and Language(Mar.2016-Mar.2020)
- ★ President-elect of the Steering Committee of ICMI (Jan. 2024-present)
- ★ Member of the Steering Committee of ICMI (Feb. 2020-present)
- ★ Program Chair for Interspeech 2026
- ★ Blue Sky Paper Chair for ACM ICMI 2023
- ★ Publication Chair of ACM ICMI 2022
- ★ General Chair of ACM ICMI 2021
- ★ Program Chair of IEEE Automatic Speech Recognition and Understanding Workshop (ASRU2021)
- ★ Workshops/Tutorials Chair of FG 2020
- ★ Workshop Chairs of ACM ICMI 2019
- ★ Publication Chair of AAAC ACII 2019
- ★ Organizer of International Workshop on Human Behavior Understanding (HBU 2018)

- ★ General Chair of AAAC ACHI 2017
- ★ Program Chair for VCIP 2017
- ★ Workshops/Tutorials Chair of FG 2017
- ★ Vice Chair of the IEEE-Dallas Chapter of SPS
- ★ Program Chair for ACM ICMI 2016
- ★ Publicity Chair of Interspeech 2016
- ★ Doctoral Spotlight Chair for ACM ICMI 2015
- ★ Workshop Chair of AAAC ACHI 2015
- ★ Doctoral Consortium Chair of IEEE FG 2015
- ★ Publicity Chair of IEEE ICME 2014
- ★ Workshop Chair of ACM ICMI 2014
- ★ Doctoral Spotlight Chair for ACM ICMI 2013
- ★ Demonstrations and Retrieval Challenges Competition Chairs at ACM ICMR 2013
- ★ Doctoral Spotlight Chair for ACM ICMI 2012
- ★ NSF review panel (CISE) (2011, 2013, 2015, 2016, 2017, 2018, 2019, 2020, 2023)
- ★ Meta Reviewer/Area Chair
 - ICASSP 2023, 2024
 - IEEE FG 2015, 2024
 - Interspeech 2015, 2016, 2019
 - ACM ICMI 2015, 2018, 2024
- ★ Technical/Program Committee
 - Generation and Evaluation of Non-verbal Behaviour for Embodied Agents (GENEA 2023)
 - International Workshop on Speech, Music, and Mind (SMM 2023)
 - Generation and Evaluation of Non-verbal Behaviour for Embodied Agents (GENEA 2022)
 - International Workshop on Human Behavior Understanding (HBU 2022)
 - Multimodal Sentiment Analysis in Real-life Media Challenge and Workshop (MuSe 2021)
 - International Workshop on Human Behavior Understanding (HBU 2021)
 - Generation and Evaluation of Non-verbal Behaviour for Embodied Agents (GENEA 2020)
 - Multimodal Sentiment Analysis in Real-life Media Challenge and Workshop (MuSe 2020)
 - Multimodal Analyses enabling Artificial Agents in Human-Machine Interaction (MA3HMI 2018)
 - Emotion Recognition in the Wild (EmotiW) Challenge at ACM ICMI 2018.
 - Speech, Music and Mind: Detecting and Influencing Mental States with Audio (2018)
 - Affective Content Analysis (AffCon2018) at AAAI-2018
 - Automatic affect analysis and synthesis at ICIAP 2017
 - Fifth Emotion Recognition in the Wild (EmotiW) challenge at ACM ICMI 2017.
 - Deep Affective Learning and Context Modeling (DAL-COM 2017)
 - International Workshop on Context Based Affect Recognition (CBAR 2016)
 - NIPS workshop on Multimodal Machine Learning 2015
 - Third Emotion Recognition in the Wild (EmotiW) Challenge at ACM ICMI 2015.
 - Second Workshop on Computer Vision for Affective Computing (CV4AC) at ICCV 2015.
 - ICME 2015

- The 3rd International Workshop on Context Based Affect Recognition (CBAR 2015)
- The 3rd International Workshop on Emotion Representation, Analysis and Synthesis in Continuous Time and Space (EmoSPACE 2015)
- ICME 2014
- ACM Multimedia 2014
- International Conference on Computational Linguistics (COLING 2014)
- Inter. Workshop on Emotion, Social Signals, Sentiment & Linked open data (ES³LOD 2014)
- ICME 2013
- International Workshop on Human Behavior Understanding (HBU 2013)
- International Workshop on Emotion Representations and Modelling for Human-Computer Interaction Systems (ERM4HCI 2013)
- The 6th Biennial Workshop on Digital Signal Processing for In-Vehicle Systems, 2013
- First Emotion Recognition In The Wild Challenge (EmotiW 2013)
- 5th International Workshop on Affective Interaction in Natural Environments (AFFINE 2013): Interacting with Affective Artefacts in the Wild
- 2nd International Workshop on Context Based Affect Recognition (CBAR 2013)
- EmoSPACE Workshop at FG 2013
- 10th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2013)
- 1st International Workshop on Context Based Affect Recognition (CBAR 2012), (SocialCom12)
- Workshop What’s in a Face?, European Conference on Computer Vision (ECCV 2012)
- 4th International Workshop on Corpora for Research on Emotion Sentiment & Social Signals (ES3 2012)
- Digital Signal Processing Workshop for In-Vehicle Systems, 2011.
- Machine Learning for Affective Computing (MLAC), 2011.
- Inferring cognitive and emotional states from multimodal behavioural measures, 2011.

SERVICE AT THE UNIVERSITY OF TEXAS AT DALLAS	Member of the EE Faculty Search Committee	2021 - present
	Member of the UTD Advisory Committee on Research	2021 - 2023
	Assigned by UTD as a mentor of junior faculty	2022–present
	Member of the UTD Library Committee	2020 - 2022
	Chair of the Ph.D. Program Committee	2020 - 2021
	Member of the Ph.D. Program Committee	2019 - 2020
	Chair of the Electrical Engineering Educational Assessment Committee	2018-2019
	Chair of the Ad-Hoc committee for the Mid-Probationary review for a colleague	2018-2019
	Member of the Electrical Engineering Educational Assessment Committee	2015-2018
	Member of the Electrical Engineering TA Assignment Committee	2012-2015
	Member of the Electrical Engineering Teaching Lab Committee	2014-2015
	Member of the Electrical Engineering Faculty Search	2012-2013
	Member of the Electrical Engineering Graduate Committee	2009-2012
	Member of the PhD committee for 84 students	2009-present

PUBLICATIONS Note: Current/former mentees are highlighted in *italics*.

Journal Articles

- [1] *S.-G. Leem*, D. Fulford, J.-P. Onnela, D. Gard, and C. Busso, “Selective acoustic feature enhancement for speech emotion recognition with noisy speech,” *IEEE/ACM Transactions on Audio, Speech and Language Processing*, vol. 32, pp. 917-929, 2024, DOI: 10.1109/TASLP.2023.3340603.

- [2] W.-C. Lin and C. Busso, "Deep temporal clustering features for speech emotion recognition," *Speech Communication*, vol. 157, pp. 103027, February 2024, DOI: 10.1016/j.specom.2023.103027
- [3] A. Vidal and C. Busso, "Multimodal attention for lip synthesis using conditional generative adversarial networks," *Speech communication*, vol. 153, pp. 102959, September 2023, DOI: 10.1016/j.specom.2023.102959.
- [4] W.-C. Lin and C. Busso, "Chunk-level speech emotion recognition: A general framework of sequence-to-one dynamic temporal modeling," *IEEE Transactions on Affective Computing*, vol. 14, no. 2, pp. 1215-1227, April-June 2023, DOI: 10.1109/TAFFC.2021.3083821
- [5] J. Harvill, S.-G. Leem, M. Abdelwahab, R. Lotfian, and C. Busso, "Quantifying emotional similarity in speech," *IEEE Transactions on Affective Computing*, vol. 14, no. 2, pp. 1376-1390, April-June 2023, DOI: 10.1109/TAFFC.2021.3127390.
- [6] Y. Qiu, T. Misu, and C. Busso, "Unsupervised scalable multimodal driving anomaly detection," *IEEE Transactions on Intelligent Vehicles*, vol. 8, no. 4, pp. 3154-3165, April 2023, DOI: 10.1109/TIV.2022.3160861.
- [7] S. Jha, N. Al-Dhahir, and C. Busso, "Driver visual attention estimation using head pose and eye appearance information," *IEEE Open Journal of Intelligent Transportation System*, vol. 4, pp. 216-231, March 2023, DOI: 10.1109/OJITS.2023.3258184.
- [8] W.-C. Lin and C. Busso, "Sequential modeling by leveraging non-uniform distribution of speech emotion," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 31, pp. 1087-1099, February 2023, DOI: 10.1109/TASLP.2023.3244527.
- [9] S. Jha and C. Busso, "Estimation of driver's gaze region from head position and orientation using probabilistic confidence regions," *IEEE Transactions on Intelligent Vehicles*, vol. 8, no. 1, pp. 59-72, January 2023. (arXiv:2012.12754). DOI: 10.1109/TIV.2022.3141071.
- [10] K. Caughlin, E. Duran-Sierra, S. Cheng, R. Cuenca, B. Ahmed, J. Ji, M. Martinez, M. Al-Khalil, H. Al-Enazi, Y.-S. L. Cheng, J. Wright, J.A. Jo, and C. Busso, "Aligning small datasets using domain adversarial learning: Applications in automated in vivo oral cancer diagnosis," *IEEE Journal of Biomedical and Health Informatics*, vol. 27, no. 1, January 2023, DOI: 10.1109/JBHI.2022.3217015.
- [11] L. Goncalves and C. Busso, "Robust audiovisual emotion recognition: Aligning modalities, capturing temporal information, and handling missing features," *IEEE Transactions on Affective Computing*, vol. 13, no. 4, pp. 2156-2170, October-December 2022, DOI: 10.1109/TAFFC.2022.3216993.
- [12] K. Sridhar and C. Busso, "Unsupervised personalization of an emotion recognition system: The unique properties of the externalization of valence in speech," *IEEE Transactions on Affective Computing*, vol. 13, no. 4, pp. 1959-1972, October-December 2022, DOI: 10.1109/TAFFC.2022.3187336.
- [13] S. Jha, M. Marzban, T. Hu, M.H. Mahmoud, N. Al-Dhahir, and C. Busso, "The multimodal driver monitoring database: A naturalistic corpus to study driver attention," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 8, pp. 10736-10752, August 2022. (arXiv:2101.04639). DOI: 10.1109/TITS.2021.3095462.
- [14] T. Hu, S. Jha, and C. Busso, "Temporal head pose estimation from point cloud in naturalistic driving conditions," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 7, pp. 8063-8076, July 2022, DOI: 10.1109/TITS.2021.3075350.
- [15] A. Vidal, S. Jha, S. Hassler, T. Price, and C. Busso, "Face detection and grimace scale prediction of white furred mice," *Machine Learning with Applications*, vol. 8, pp. 100312, June 2022, DOI: 10.1016/j.mlwa.2022.100312.
- [16] N. Sadoughi and C. Busso, "Speech-driven expressive talking lips with conditional sequential generative adversarial networks," *IEEE Transactions on Affective Computing*, vol. 12, no. 4, pp. 1031-1044, October-December 2021. (ArXiv: 1806.00154). DOI: 10.1109/TAFFC.2019.2916031.
- [17] R. Lotfian and C. Busso, "Over-sampling emotional speech data based on subjective evaluations provided by multiple individuals," *IEEE Transactions on Affective Computing*, vol. 4, no. 12, pp. 870-882, October-December 2021, DOI: 10.1109/TAFFC.2019.2901465.
- [18] C.-C. Lee, K. Sridhar, J.-L. Li, W.-C. Lin, B.-H. Su, and C. Busso, "Deep Representation Learning for Affective Speech Signal Analysis and Processing: Preventing unwanted signal disparities,"

- IEEE Signal Processing Magazine, vol. 38, no. 6, pp. 22-38, November 2021, DOI: 10.1109/MSP.2021.3105939.
- [19] E. Duran-Sierra, S. Cheng, R. Cuenca, B. Ahmed, J. Ji, V. Yakovlev, M. Martinez, M. Al-Khalil, H. Al-Enazi, Y.S.L. Cheng, J. Wright, C. Busso, and J. Jo, "Machine-learning assisted discrimination of precancerous and cancerous from healthy oral tissue based on multispectral autofluorescence lifetime imaging endoscopy," *Cancers*, vol. 13, no. 19, pp. 1-16, September 2021, DOI: 10.3390/cancers13194751.
 - [20] K. N. Haque, R. Rana, J. Liu, J.H.L. Hansen, N. Cummins, C. Busso, and B. Schuller, "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*, vol. 29, pp. 2575-2590, July 2021, DOI: 10.1109/TASLP.2021.3098764.
 - [21] D. Fulford, J. Mote, R. Gonzalez, S. Abplanalp, Y. Zhang, *J. Luckenbaugh*, J.-P. Onnela, C. Busso, and D. Gard, "Smartphone sensing of social interactions in people with and without schizophrenia," *Journal of Psychiatric Research*, vol. 137, pp. 613-620, May 2021, DOI: 10.1016/j.jpsychires.2020.11.002.
 - [22] *S. Parthasarathy* and C. Busso, "Predicting emotionally salient regions using qualitative agreement of deep neural network regressors," *IEEE Transactions on Affective Computing*, vol. 12, no. 2, pp. 402-416, April-June 2021, DOI: 10.1109/TAFFC.2018.2878715.
 - [23] G.N. Yannakakis, R. Cowie, and C. Busso, "The ordinal nature of emotions: An emerging approach," *IEEE Transactions on Affective Computing*, vol. 12, no. 1, pp. 16-35, January-March 2021, DOI: 10.1109/TAFFC.2018.2879512. **Best Paper Award from IEEE Transactions on Affective Computing.**
 - [24] *F. Tao* and C. Busso, "End-to-end audiovisual speech recognition system with multi-task learning," *IEEE Transactions on Multimedia*, vol. 23, pp. 1-11, January 2021, DOI: 10.1109/TMM.2020.2975922.
 - [25] *S. Parthasarathy* and C. Busso, "Semi-supervised speech emotion recognition with ladder networks," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 28, pp. 2697-2709, September 2020. (ArXiv: 1905.02921). DOI: 10.1109/TASLP.2020.3023632.
 - [26] *R. Lotfian* and C. Busso, "Building naturalistic emotionally balanced speech corpus by retrieving emotional speech from existing podcast recordings," *IEEE Transactions on Affective Computing*, vol. 10, no. 4, pp. 471-483, October-December 2019, DOI: 10.1109/TAFFC.2017.2736999. **Best of IEEE Transactions on Affective Computing Paper Collection award.**
 - [27] *A. Vidal*, J. Silva, and C. Busso, "Discriminative features for texture retrieval using wavelet packets," *IEEE Access*, vol. 7, no. 1, pp. 148882-148896, December 2019, DOI: 10.1109/ACCESS.2019.2947006.
 - [28] *R. Lotfian* and C. Busso, "Lexical dependent emotion detection using synthetic speech reference," *IEEE Access*, vol. 7, no. 1, pp. 22071-22085, December 2019, DOI: 10.1109/ACCESS.2019.2898353.
 - [29] *F. Tao* and C. Busso, "End-to-end audiovisual speech activity detection with bimodal recurrent neural models," *Speech Communication*, vol. 113, pp. 25-35, October 2019. (arXiv:1809.04553), DOI: 10.1016/j.specom.2019.07.003.
 - [30] *N. Sadoughi* and C. Busso, "Speech-driven animation with meaningful behaviors," *Speech Communication*, vol. 110, pp. 90-100, July 2019. (ArXiv: 1708.01640). DOI: 10.1016/j.specom.2019.04.005.
 - [31] *R. Lotfian* and C. Busso, "Curriculum learning for speech emotion recognition from crowdsourced labels," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 27, no. 4, pp. 815-826, April 2019. (ArXiv: 1805.10339). DOI: 10.1109/TASLP.2019.2898816.
 - [32] *M. Abdelwahab* and C. Busso, "Domain adversarial for acoustic emotion recognition," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 26, no. 12, pp. 2423-2435, December 2018. (ArXiv: 1804.07690). DOI: 10.1109/TASLP.2018.2867099.
 - [33] *F. Tao* and C. Busso, "Gating neural network for large vocabulary audiovisual speech recognition," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 26, no. 7, pp. 1286-1298, July 2018, DOI: 10.1109/TASLP.2018.2815268.
 - [34] *N. Li* and C. Busso, "Calibration free, user independent gaze estimation with tensor analysis,"

- Image and Vision Computing, vol. 74, pp. 10-20, June 2018, DOI: 10.1016/j.imavis.2018.04.001.
- [35] *N. Sadoughi*, Y. Liu, and C. Busso, "Meaningful head movements driven by emotional synthetic speech," *Speech Communication*, vol. 95, pp. 87-99, December 2017, DOI: 10.1016/j.specom.2017.07.004.
 - [36] J.H.L. Hansen, C. Busso, Y. Zheng, and A. Sathyanarayana, "Driver modeling for detection and assessment of driver distraction: Examples from the UTDriVe test bed," *IEEE Signal Processing Magazine*, vol. 34, no. 4, pp. 130-142, July 2017, DOI: 10.1109/MSP.2017.2699039.
 - [37] *S. Mariooryad* and C. Busso, "The cost of dichotomizing continuous labels for binary classification problems: Deriving a Bayesian-optimal classifier," *IEEE Transactions on Affective Computing*, vol. 8, no. 1, pp. 67-80 January-March 2017, DOI: 10.1109/TAFFC.2015.2508454.
 - [38] C. Busso, *S. Parthasarathy*, *A. Burmania*, *M. AbdelWahab*, *N. Sadoughi*, and E. Mower Provost, "MSP-IMPROV: An acted corpus of dyadic interactions to study emotion perception," *IEEE Transactions on Affective Computing*, vol. 8, no. 1, pp. 119-130 January-March 2017, DOI: 10.1109/TAFFC.2016.2515617.
 - [39] *S. Parthasarathy*, R. Cowie, and C. Busso, "Using agreement on direction of change to build rank-based emotion classifiers," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 24, no. 11, pp. 2108-2121, November 2016, DOI: 10.1109/TASLP.2016.2593944.
 - [40] *N. Li* and C. Busso, "Detecting drivers' mirror-checking actions and its application to maneuver and secondary task recognition," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 4, pp. 980-992, April 2016, DOI: 10.1109/TITS.2015.2493451.
 - [41] *A. Burmania*, *S. Parthasarathy*, and C. Busso, "Increasing the reliability of crowdsourcing evaluations using online quality assessment," *IEEE Transactions on Affective Computing*, vol. 7, no. 4, pp. 374-388, October-December 2016, DOI: 10.1109/TAFFC.2015.2493525.
 - [42] *S. Mariooryad* and C. Busso, "Facial expression recognition in the presence of speech using blind lexical compensation," *IEEE Transactions on Affective Computing*, vol. 7, no. 4, pp. 346-359, October-December 2016, DOI: 10.1109/TAFFC.2015.2490070.
 - [43] F. Eyben, K. Scherer, B. Schuller, J. Sundberg, E. André, C. Busso, L. Devillers, J. Epps, P. Laukka, S. Narayanan, and K. Truong, "The Geneva minimalistic acoustic parameter set (GeMAPS) for voice research and affective computing," *IEEE Transactions on Affective Computing*, vol. 7, no. 2, pp. 190-202, April-June 2016, DOI: 10.1109/TAFFC.2015.2457417.
 - [44] A. Metallinou, Z. Yang, C.-C. Lee, C. Busso, S. Carnicke, and S. Narayanan, "The USC CreativeIT database of multimodal dyadic interactions: From speech and full body motion capture to continuous emotional annotations," *Journal of Language Resources and Evaluation*, vol. 50, no. 3, pp. 497-521, September 2016, DOI: 10.1007/s10579-015-9300-0.
 - [45] E. Mower Provost, Y. Shangguan, and C. Busso, "UMEME: University of Michigan emotional McGurk effect data set," *IEEE Transactions on Affective Computing*, vol. 6, no. 4, pp. 395-409, October-December 2015, DOI: 10.1109/TAFFC.2015.2407898.
 - [46] C. Poellabauer, N. Yadav, L. Daudet, S. Schneider, C. Busso, and P. Flynn, "Challenges in concussion detection using vocal acoustic biomarkers," *IEEE Access*, vol. 3, pp. 1143-1160, August 2015, DOI: 10.1109/ACCESS.2015.2457392. **IEEE Access Best Multimedia Contest Winner.**
 - [47] *S. Mariooryad* and C. Busso, "Correcting time-continuous emotional labels by modeling the reaction lag of evaluators," *IEEE Transactions on Affective Computing*, vol. 6, no. 2, pp. 97-108, April-June 2015, DOI: 10.1109/TAFFC.2014.2334294. **Special Issue Best of ACII 2013.**
 - [48] *N. Li* and C. Busso, "Predicting perceived visual and cognitive distractions of drivers with multimodal features," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 1, pp. 51-65, February 2015, DOI: 10.1109/TITS.2014.2324414.
 - [49] *S. Mariooryad* and C. Busso, "Compensating for speaker or lexical variabilities in speech for emotion recognition," *Speech Communication*, vol. 57, pp. 1-12, February 2014, DOI: 10.1016/j.specom.2013.07.011.
 - [50] *J.P. Arias*, C. Busso, and N.B. Yoma, "Shape-based modeling of the fundamental frequency contour for emotion detection in speech," *Computer Speech and Language*, vol. 28, no. 1, pp. 278-294, January 2014, DOI: 10.1016/j.csl.2013.07.002.
 - [51] C. Busso, *S. Mariooryad*, A. Metallinou, and S. Narayanan, "Iterative feature normalization

- scheme for automatic emotion detection from speech,” *IEEE Transactions on Affective Computing*, vol. 4, no. 4, pp. 386-397, October-December 2013, DOI: 10.1109/T-AFFC.2013.26.
- [52] *S. Mariooryad* and C. Busso, “Exploring cross-modality affective reactions for audiovisual emotion recognition,” *IEEE Transactions on Affective Computing*, vol. 4, no. 2, pp. 183-196, April-June 2013, DOI: 10.1109/T-AFFC.2013.11.
 - [53] *N. Li, J.J. Jain*, and C. Busso, “Modeling of driver behavior in real world scenarios using multiple noninvasive sensors,” *IEEE Transactions on Multimedia*, vol. 15, no. 5, pp. 1213-1225, August 2013, DOI: 10.1109/TMM.2013.2241416.
 - [54] *S. Mariooryad* and C. Busso, “Generating human-like behaviors using joint, speech-driven models for conversational agents,” *IEEE Transactions on Audio, Speech and Language Processing*, vol. 20, no. 8, pp. 2329-2340, October 2012, DOI: 10.1109/TASL.2012.2201476.
 - [55] C.-C. Lee, E. Mower, C. Busso, S. Lee, and S. Narayanan, “Emotion Recognition Using a Hierarchical Binary Decision Tree Approach,” *Speech Communication*, vol. 53, no. 9-10, pp. 1162-1171, November-December 2011, DOI: 10.1016/j.specom.2011.06.004.
 - [56] C. Busso, S. Lee, and S.S. Narayanan, “Analysis of emotionally salient aspects of fundamental frequency for emotion detection,” *IEEE Transactions on Audio, Speech and Language Processing*, vol. 17, no. 4, pp. 582-596, May 2009, DOI: 10.1109/TASL.2008.2009578.
 - [57] C. Busso, M. Bulut, C.C. Lee, A. Kazemzadeh, E. Mower, S. Kim, J.N. Chang, S. Lee, and S.S. Narayanan, “IEMOCAP: Interactive emotional dyadic motion capture database,” *Journal of Language Resources and Evaluation*, vol. 42, no. 4, pp. 335-359, December 2008, DOI: 10.1007/s10579-008-9076-6.
 - [58] C. Busso and S. Narayanan, “Interrelation between speech and facial gestures in emotional utterances: a single subject study,” *IEEE Transactions on Audio, Speech and Language Processing*, vol. 15, no. 8, pp. 2331-2347, November 2007, DOI: 10.1109/TASL.2007.905145.
 - [59] C. Busso, Z. Deng, M. Grimm, U. Neumann, and S. Narayanan, “Rigid head motion in expressive speech animation: Analysis and synthesis,” *IEEE Transactions on Audio, Speech and Language Processing*, vol. 15, no. 3, pp. 1075-1086, March 2007, DOI: 10.1109/TASL.2006.885910.
 - [60] C. Busso, Z. Deng, U. Neumann, and S.S. Narayanan, “Natural head motion synthesis driven by acoustic prosodic features,” *Computer Animation and Virtual Worlds*, vol. 16, no. 3-4, pp. 283-290, July 2005, DOI: 10.1002/cav.80.
 - [61] N.B. Yoma, C. Molina, J. Silva, and C. Busso, “Modeling, estimating, and compensating low-bit rate coding distortion in speech recognition,” *IEEE Transactions on Audio, Speech and Language Processing*, vol. 14, no. 1, pp. 246-255, January 2006, DOI: 10.1109/TSA.2005.852994.
 - [62] N.B. Yoma, C. Busso, and I. Soto, “Packet-loss modelling in IP networks with state-duration constraints,” *Communications, IEE Proceedings*, vol. 152, no. 1, pp. 1-5, Feb 2005, DOI: 10.1049/ip-com:20040774.
 - [63] N.B. Yoma, J. Hood, and C. Busso, “A real-time protocol for the internet based on the least mean square algorithm,” *Transactions on Multimedia, IEEE*, vol. 6, no. 1, pp. 174-184, Feb 2004, DOI: 10.1109/TMM.2003.819582.
 - [64] N.B. Yoma, J. Silva, C. Busso, and I. Brito, “Compensating additive noise and CS-CELP distortion in speech recognition using stochastic weighted Viterbi algorithm,” *Electronics Letters, IEE*, vol. 39, no. 4, pp. 409-411, Feb 2003, DOI: 10.1049/el:20030252.

Conference Proceedings

- [1] *L. Martinez-Lucas* and C. Busso, “Dynamic speech emotion recognition using a conditional neural process,” in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2024)*, Seoul, Republic of Korea, April 2024, vol. To appear.
- [2] *A. Reddy Naini*, M.A. Kohler, E. Richerson, D. Robinson, and C. Busso, “Generalization of self-supervised learning-based representations for cross-domain speech emotion recognition,” in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2024)*, Seoul, Republic of Korea, April 2024, vol. To appear.
- [3] I.R. Ulgen, Z. Du, C. Busso, and B. Sisman, “Revealing emotional clusters in speaker embeddings: A contrastive learning strategy for speech emotion recognition,” in *IEEE International Conference*

on Acoustics, Speech, and Signal Processing (ICASSP 2024), Seoul, Republic of Korea, April 2024, vol. To appear.

- [4] *A. Reddy Naini, S. Subramaniam, S.-G. Leem*, and C. Busso, “Combining relative and absolute learning formulations to predict emotional attributes from speech,” in *IEEE Automatic Speech Recognition and Understanding Workshop (ASRU 2023)*, Taipei, Taiwan, December 2023.
- [5] *W.-C. Lin, L. Goncalves*, and C. Busso, “Enhancing resilience to missing data in audio-text emotion recognition with multi-scale chunk regularization,” in *ACM International Conference on Multimodal Interaction (ICMI 2023)*, Paris, France, October 2023, pp. 207-215, DOI: 10.1145/3577190.3614110.
- [6] *I. Brooks, S. Gogineni, S. Jha*, S.J. Ray, R. Narasimha, N. Al-Dhahir, and C. Busso, “MSP-DISK: Naturalistic and diverse in-vehicle database for joint pose and seat belt detection,” in *IEEE International Conference on Intelligent Transportation Systems (ITSC 2023)*, Bilbao, Spain, September 2023.
- [7] *L. Martinez-Lucas, A. Salman, S.-G. Leem*, S.G. Upadhyay, C.-C. Lee, and C. Busso, “Analyzing the effect of affective priming on emotional annotations,” in *International Conference on Affective Computing and Intelligent Interaction (ACII 2023)*, Cambridge, MA, USA, September 2023.
- [8] S.G. Upadhyay, W.-S. Chien, B.-H. Su, *L. Goncalves*, Y.-T. Wu, *A.N. Salman*, C. Busso, and C.-C. Lee, “An intelligent infrastructure toward large scale naturalistic affective speech corpora collection,” in *International Conference on Affective Computing and Intelligent Interaction (ACII 2023)*, Cambridge, MA, USA, September 2023.
- [9] *S.-G. Leem*, D. Fulford, J.-P. Onnela, D. Gard, and C. Busso, “Computation and memory efficient noise adaptation of Wav2Vec2.0 for noisy speech emotion recognition with skip connection adapters,” in *Interspeech 2023*, Dublin, Ireland, August 2023, pp. 1888-1892, DOI: 10.21437/Interspeech.2023-1034.
- [10] *A. Reddy Naini, A. Salman*, and C. Busso, “Preference learning labels by anchoring on consecutive annotations,” in *Interspeech 2023*, Dublin, Ireland, August 2023, pp. 1898-1902, DOI: 10.21437/Interspeech.2023-1108.
- [11] N. Grágeda, C. Busso, E. Alvarado, R. Mahu, and N. Becerra Yoma, “Distant speech emotion recognition in an indoor human-robot interaction scenario,” in *Interspeech 2023*, Dublin, Ireland, August 2023, pp. 3657-3661, DOI: 10.21437/Interspeech.2023-1169.
- [12] *H.-C. Chou, L. Goncalves, S.-G. Leem*, C.-C. Lee, and C. Busso, “The importance of calibration: Rethinking confidence and performance of speech multi-label emotion classifiers,” in *Interspeech 2023*, Dublin, Ireland, August 2023, pp. 641-645. DOI: 10.21437/Interspeech.2023-1113.
- [13] *S. Jha, I. Brooks*, S.J. Ray, R. Narasimha, N. Al-Dhahir, and C. Busso, “Seatbelt segmentation using synthetic images,” in *IEEE Intelligent Vehicles Symposium (IV 2023)*, Anchorage, AK, USA, June 2023, pp. 1-6, DOI: 10.1109/IV55152.2023.10186571.
- [14] *Y. Qiu*, T. Misu, and C. Busso, “Example-based query to identify cause of driving anomaly with few labeled samples,” in *IEEE Intelligent Vehicles Symposium (IV 2023)*, Anchorage, AK, USA, June 2023, pp. 1-7, DOI: 10.1109/IV55152.2023.10186733.
- [15] *S.-G. Leem*, D. Fulford, J.-P. Onnela, D. Gard, and C. Busso, “Adapting a self-supervised speech representation for noisy speech emotion recognition by using contrastive teacher-student learning,” in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2023)*, Rhodes island, Greece, 2023, pp. 1-5, DOI: 10.1109/ICASSP49357.2023.10097135.
- [16] *W.-C. Lin* and C. Busso, “Role of lexical boundary information in chunk-level segmentation for speech emotion recognition,” in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2023)*, Rhodes island, Greece, 2023, pp. 1-5, DOI: 10.1109/ICASSP49357.2023.10096861.
- [17] *L. Goncalves* and C. Busso, “Learning cross-modal audiovisual representations with ladder networks for emotion recognition,” in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2023)*, Rhodes island, Greece, 2023, pp. 1-5, DOI: 10.1109/ICASSP49357.2023.10096138.
- [18] *A. Reddy Naini*, M.A. Kohler, and C. Busso, “Unsupervised domain adaptation for preference learning based speech emotion recognition,” in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2023)*, Rhodes island, Greece, 2023, pp. 1-5,

DOI: 10.1109/ICASSP49357.2023.10094301.

- [19] S.G. Upadhyay, *L. Martinez-Lucas*, B.-H. Su, *W.-C. Lin*, W.-S. Chien, Y.-T. Wu, W. Katz, C. Busso, and C.-C. Lee, "Phonetic anchor-based transfer learning to facilitate unsupervised cross-lingual speech emotion recognition," in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2023), Rhodes island, Greece, 2023, pp. 1-5, DOI: 10.1109/ICASSP49357.2023.10095250.
- [20] *A. N. Salman* and C. Busso, "Privacy preserving personalization for video facial expression recognition using federated learning," in ACM International Conference on Multimodal Interaction (ICMI 2022), Bangalore, India, November 2022, pp. 495-503, DOI: 10.1145/3536221.3556614.
- [21] W.-S. Chien, S. Upadhyay, *W.-C. Lin*, Y.-T. Wu, B.-H. Su, C. Busso, and C.-C. Lee, "Monologue versus conversation: Differences in emotion perception and acoustic expressivity," in International Conference on Affective Computing and Intelligent Interaction (ACII 2022), Nara, Japan, October 2022, pp. 1-7, DOI: 10.1109/ACII55700.2022.9953814.
- [22] Y. Qiu, T. Misu, and C. Busso, "Driving anomaly detection using contrastive multiview coding to interpret cause of anomaly," in IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2022), Kyoto, Japan, October 2022, pp. 11424-11431, DOI: 10.1109/IROS47612.2022.9981815.
- [23] *L. Goncalves* and C. Busso, "Improving speech emotion recognition using self-supervised learning with domain-specific audiovisual tasks," in Interspeech 2022, Incheon, South Korea, September 2022, pp. 1168-1172, DOI: 10.21437/Interspeech.2022-11012.
- [24] *H.-C. Chou*, C.-C. Lee, and C. Busso, "Exploiting co-occurrence frequency of emotions in perceptual evaluations to train a speech emotion classifier," in Interspeech 2022, Incheon, South Korea, September 2022, pp. 161-165, DOI: 10.21437/Interspeech.2022-11041.
- [25] *L. Goncalves* and C. Busso, "AuxFormer: Robust approach to audiovisual emotion recognition," in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2022), Singapore, May 2022, pp. 7357-7361, DOI: 10.1109/ICASSP43922.2022.9747157.
- [26] *S.-G. Leem*, D. Fulford, J.-P. Onnela, D. Gard, and C. Busso, "Not all features are equal: Selection of robust features for speech emotion recognition in noisy environments," in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2022), Singapore, May 2022, pp. 6447-6451, DOI: 10.1109/ICASSP43922.2022.9747705.
- [27] Y. Qiu, C. Busso, T. Misu, and K. Akash, "Incorporating gaze behavior using joint embedding with scene context for driver takeover detection," in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2022), Singapore, May 2022, pp. 4633-4637, DOI: 10.1109/ICASSP43922.2022.9747779.
- [28] *H.-C. Chou*, *W.-C. Lin*, C.-C. Lee, and C. Busso, "Exploiting annotators' typed description of emotion perception to maximize utilization of ratings for speech emotion recognition," in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2022), Singapore, May 2022, pp. 7717-7721, DOI: 10.1109/ICASSP43922.2022.9746990.
- [29] *K. Caughlin*, E. Duran-Sierra, S. Cheng, R. Cuenca, B. Ahmed, J. Ji, V.V. Yakovlev, M. Martinez, M. Al-Khalil, H. Al-Enazi, J. A. Jo, and C. Busso, "End-to-end neural network for feature extraction and cancer diagnosis of in vivo fluorescence lifetime images of oral lesions," in International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC 2021), Guadalajara, Mexico, November 2021, pp. 3894-3897, DOI: 10.1109/EMBC46164.2021.9629739.
- [30] *K. Sridhar*, *W.-C. Lin*, and C. Busso, "Generative approach using soft-labels to learn uncertainty in predicting emotional attributes," in International Conference on Affective Computing and Intelligent Interaction (ACII 2021), Nara, Japan, September-October 2021, pp. 1-8, DOI: 10.1109/ACII52823.2021.9597461.
- [31] *S.-G. Leem*, D. Fulford, J.-P. Onnela, D. Gard, and C. Busso, "Separation of emotional and reconstruction embeddings on ladder network to improve speech emotion recognition robustness in noisy conditions," in Interspeech 2021, Brno, Czech Republic, August-September 2021, pp. 2871-2875, DOI: 10.21437/Interspeech.2021-1438.
- [32] *J. Luckenbaugh*, S. Abplanalp, R. Gonzalez, D. Fulford, D. Gard, and C. Busso, "Voice activity detection with teacher-student domain emulation," in Interspeech 2021, Brno, Czech Republic, August-September 2021, pp. 4374-4378, DOI: 10.21437/Interspeech.2021-1234.

- [33] *W.-C. Lin, K. Sridhar, and C. Busso*, “DeepEmoCluster: A semi-supervised framework for latent cluster representation of speech emotions,” in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2021)*, Toronto, ON, Canada, June 2021, pp. 7263-7267, DOI: 10.1109/ICASSP39728.2021.9414035.
- [34] *A. Vidal, A. Salman, W.-C. Lin, and C. Busso*, “MSP-face corpus: A natural audiovisual emotional database,” in *ACM International Conference on Multimodal Interaction (ICMI 2020)*, Utrecht, The Netherlands, October 2020, pp. 397-405, DOI: 10.1145/3382507.3418872.
- [35] *W.-C. Lin and C. Busso*, “An efficient temporal modeling approach for speech emotion recognition,” in *Interspeech 2020*, Shanghai, China, October 2020, pp. 2322-2326, DOI: 10.21437/Interspeech.2020-2636.
- [36] *K. Sridhar and C. Busso*, “Ensemble of students taught by probabilistic teachers to improve speech emotion recognition,” in *Interspeech 2020*, Shanghai, China, October 2020, pp. 516-520, DOI: 10.21437/Interspeech.2020-2694.
- [37] *L. Martinez-Lucas, M. Abdelwahab, and C. Busso*, “The MSP-conversation corpus,” in *Interspeech 2020*, Shanghai, China, October 2020, pp. 1823-1827, DOI: 10.21437/Interspeech.2020-2444.
- [38] *A. N. Salman and C. Busso*, “Style extractor for facial expression recognition in the presence of speech,” in *IEEE International Conference on Image Processing (ICIP 2020)*, Abu Dhabi, United Arab Emirates (UAE), October 2020, pp. 1806-1810, DOI: 10.1109/ICIP40778.2020.9191330.
- [39] *Y. Qiu, T. Misu, and C. Busso*, “Use of triplet loss function to improve driving anomaly detection using conditional generative adversarial network,” in *Intelligent Transportation Systems Conference (ITSC 2020)*, Rhodes, Greece, September 2020, pp. 1-7, DOI: 10.1109/ITSC45102.2020.9294562.
- [40] *T. Hu, S. Jha, and C. Busso*, “Robust driver head pose estimation in naturalistic conditions from point-cloud data,” in *IEEE Intelligent Vehicles Symposium (IV2020)*, Las Vegas, NV USA, October-November 2020, pp. 1176-1182, DOI: 10.1109/IV47402.2020.9304592.
- [41] *A.N. Salman and C. Busso*, “Dynamic versus static facial expressions in the presence of speech,” in *IEEE International Conference on Automatic Face and Gesture Recognition (FG 2020)*, Buenos Aires, Argentina, May 2020, pp. 436-443, DOI: 10.1109/FG47880.2020.00119.
- [42] *K. Sridhar and C. Busso*, “Modeling uncertainty in predicting emotional attributes from spontaneous speech” in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2020)*, Barcelona, Spain, May 2020, pp. 8384-8388, DOI: 10.1109/ICASSP40776.2020.9054237.
- [43] *Y. Qiu, T. Misu, and C. Busso*, “Analysis of the relationship between physiological signals and vehicle maneuvers during a naturalistic driving study,” in *Intelligent Transportation Systems Conference (ITSC 2019)*, Auckland, New Zealand, October 2019, pp. 3230-3235, DOI: 10.1109/ITSC.2019.8917198.
- [44] *Y. Qiu, T. Misu, and C. Busso*, “Driving anomaly detection with conditional generative adversarial network using physiological and can-bus data,” in *ACM International Conference on Multimodal Interaction (ICMI 2019)*, Suzhou, Jiangsu, China, October 2019, pp. 164-173, DOI: 10.1145/3340555.3353749.
- [45] *K. Sridhar and C. Busso*, “Speech emotion recognition with a reject option,” in *Interspeech 2019*, Graz, Austria, September 2019, pp. 3272-3276, DOI: 10.21437/Interspeech.2019-1842.
- [46] *M. Abdelwahab and C. Busso*, “Active learning for speech emotion recognition using deep neural network,” in *International Conference on Affective Computing and Intelligent Interaction (ACII 2019)*, Cambridge, UK, September 2019, pp. 441-447, DOI: 10.1109/ACII.2019.8925524.
- [47] *M. Bancroft, R. Lotfian, J. Hansen, and C. Busso*, “Exploring the intersection between speaker verification and emotion recognition,” in *International Workshop on Social & Emotion AI for Industry (SEAIxI)*, Cambridge, UK, September 2019, pp. 337-342, DOI: 10.1109/ACIIW.2019.8925044.
- [48] *J. Harvill, M. AbdelWahab, R. Lotfian, and C. Busso*, “Retrieving speech samples with similar emotional content using a triplet loss function,” in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2019)*, Brighton, UK, May 2019, pp. 7400-7404, DOI: 10.1109/ICASSP.2019.8683273.
- [49] *S. Jha and C. Busso*, “Estimation of gaze region using two dimensional probabilistic maps con-

- structed using convolutional neural networks,” in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2019), Brighton, UK, May 2019, pp. 3792-3796, DOI: 10.1109/ICASSP.2019.8683794.
- [50] *S. Jha* and C. Busso, “Probabilistic estimation of the gaze region of the driver using dense classification,” in IEEE International Conference on Intelligent Transportation (ITSC 2018), Maui, HI, USA, November 2018, pp. 697-702, DOI: 10.1109/ITSC.2018.8569709.
 - [51] *S. Parthasarathy* and C. Busso, “Ladder networks for emotion recognition: Using unsupervised auxiliary tasks to improve predictions of emotional attributes,” in Interspeech 2018, Hyderabad, India, September 2018, pp. 3698-3702. (ArXiv:1804.10816), DOI: 10.21437/Interspeech.2018-1391.
 - [52] *F. Tao* and C. Busso, “Audiovisual speech activity detection with advanced long short-term memory,” in Interspeech 2018, Hyderabad, India, September 2018, pp. 1244-1248, DOI: 10.21437/Interspeech.2018-2490.
 - [53] *K. Sridhar*, *S. Parthasarathy*, and C. Busso, “Role of regularization in the prediction of valence from speech,” in Interspeech 2018, Hyderabad, India, September 2018, pp. 941-945, DOI: 10.21437/Interspeech.2018-2508.
 - [54] *S. Parthasarathy* and C. Busso, “Preference-learning with qualitative agreement for sentence level emotional annotations,” in Interspeech 2018, Hyderabad, India, September 2018, pp. 252-256, DOI: 10.21437/Interspeech.2018-2478.
 - [55] *R. Lotfian* and C. Busso, “Predicting categorical emotions by jointly learning primary and secondary emotions through multitask learning,” in Interspeech 2018, Hyderabad, India, September 2018, pp. 951-955, DOI: 10.21437/Interspeech.2018-2464.
 - [56] *F. Tao* and C. Busso, “Aligning audiovisual features for audiovisual speech recognition,” in IEEE International Conference on Multimedia and Expo (ICME 2018), San Diego, CA, USA, July 2018, pp. 1-6, DOI: 10.1109/ICME.2018.8486455.
 - [57] *S. Jha* and C. Busso, “Fi-Cap: Robust framework to benchmark head pose estimation in challenging environments,” in IEEE International Conference on Multimedia and Expo (ICME 2018), San Diego, CA, USA, July 2018, pp. 1-6, DOI: 10.1109/ICME.2018.8486490.
 - [58] *N. Sadoughi* and C. Busso, “Expressive speech-driven lip movements with multitask learning,” in IEEE Conference on Automatic Face and Gesture Recognition (FG 2018), Xi’an, China, May 2018, pp. 409-415, DOI: 10.1109/FG.2018.00066.
 - [59] *M. Abdelwahab* and C. Busso, “Study of dense network approaches for speech emotion recognition,” in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2018), Calgary, AB, Canada, April 2018, pp. 5084-5088, DOI: 10.1109/ICASSP.2018.8461866.
 - [60] *N. Sadoughi* and C. Busso, “Novel realizations of speech-driven head movements with generative adversarial networks,” in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2018), Calgary, AB, Canada, April 2018, pp. 6169-6173, DOI: 10.1109/ICASSP.2018.8461967.
 - [61] *M. Bokshi*, *F. Tao*, C. Busso, and J.H. L. Hansen, “Assessment and classification of singing quality based on audio-visual features,” in IEEE Visual Communications and Image Processing (VCIP 2017), St. Petersburg, FL, USA, December 2017, pp. 1-4, DOI: 10.1109/VCIP.2017.8305078.
 - [62] *R. Lotfian* and C. Busso, “Formulating emotion perception as a probabilistic model with application to categorical emotion classification,” in International Conference on Affective Computing and Intelligent Interaction (ACII 2017), San Antonio, TX, USA, October 2017, pp. 415-420, DOI: 10.1109/ACII.2017.8273633.
 - [63] *S. Parthasarathy* and C. Busso, “Predicting speaker recognition reliability by considering emotional content,” in International Conference on Affective Computing and Intelligent Interaction (ACII 2017), San Antonio, TX, USA, October 2017, pp. 434-436, DOI: 10.1109/ACII.2017.8273636.
 - [64] *G.N. Yannakakis*, *R. Cowie*, and C. Busso, “The ordinal nature of emotions,” in International Conference on Affective Computing and Intelligent Interaction (ACII 2017), San Antonio, TX, USA, October 2017, pp. 248-255, DOI: 10.1109/ACII.2017.8273608. **AAAC ACII 2017 Best Paper Award**
 - [65] *S. Jha* and C. Busso, “Probabilistic estimation of the driver’s gaze from head orientation and position,” in IEEE International Conference on Intelligent Transportation (ITSC), Yokohama,

- Japan, October 2017, pp. 1630-1635, DOI: 10.1109/ITSC.2017.8317841.
- [66] *S. Jha* and C. Busso, "Challenges in head pose estimation of drivers in naturalistic recordings using existing tools," in IEEE International Conference on Intelligent Transportation (ITSC), Yokohama, Japan, October 2017, pp. 1624-1629, DOI: 10.1109/ITSC.2017.8317870.
 - [67] *N. Sadoughi* and C. Busso, "Joint learning of speech-driven facial motion with bidirectional long-short term memory," in International Conference on Intelligent Virtual Agents (IVA 2017), J. Beskow, C. Peters, G. Castellano, C. O'Sullivan, I. Leite, S. Kopp, Eds., vol. 10498 of Lecture Notes in Computer Science, pp. 389-402. Springer Berlin Heidelberg, Stockholm, Sweden, August 2017, DOI: 10.1007/978-3-319-67401-8 49.
 - [68] *F. Tao* and C. Busso, "Bimodal recurrent neural network for audiovisual voice activity detection," in Interspeech 2017, Stockholm, Sweden, August 2017, pp. 1938-1942, DOI: 10.21437/Interspeech.2017-1573.
 - [69] *S. Parthasarathy* and C. Busso, "Jointly predicting arousal, valence and dominance with multi-task learning," in Interspeech 2017, Stockholm, Sweden, August 2017, pp. 1103-1107, DOI: 10.21437/Interspeech.2017-1494. **Nominated for Best Student Paper at Interspeech 2017.**
 - [70] *A. Burmania* and C. Busso, "A stepwise analysis of aggregated crowdsourced labels describing multimodal emotional behaviors," in Interspeech 2017, Stockholm, Sweden, August 2017, pp. 152-157, DOI: 10.21437/Interspeech.2017-1278.
 - [71] *M. Abdelwahab* and C. Busso, "Incremental adaptation using active learning for acoustic emotion recognition," in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2017), New Orleans, LA, USA, March 2017, pp. 5160-5164, DOI: 10.1109/ICASSP.2017.7953140.
 - [72] *S. Parthasarathy* and *R. Lotfian* and C. Busso, "Ranking emotional attributes with deep neural networks," in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2017), New Orleans, LA, USA, March 2017, pp. 4995-4999, DOI: 10.1109/ICASSP.2017.7953107.
 - [73] *M. Abdelwahab* and C. Busso, "Ensemble feature selection for domain adaptation in speech emotion recognition," in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2017), New Orleans, LA, USA, March 2017, pp. 5000-5004, DOI: 10.1109/ICASSP.2017.7953108.
 - [74] *S. Parthasarathy*, C. Zhang, J.H.L. Hansen, and C. Busso, "A study of speaker verification performance with expressive speech," in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2017), New Orleans, LA, USA, March 2017, pp. 5540-5544, DOI: 10.1109/ICASSP.2017.7953216.
 - [75] *S. Jha* and C. Busso, "Analyzing the Relationship Between Head Pose and Gaze to Model Driver Visual Attention", in IEEE Conference on Intelligent Transportation Systems (ITSC 2016), Rio de Janeiro, Brazil, November 2016, pp. 2157-2162, DOI: 10.1109/ITSC.2016.7795905.
 - [76] *N. Sadoughi* and C. Busso, "Head motion generation with synthetic speech: a data driven approach," in Interspeech 2016, San Francisco, CA, USA, September 2016, pp. 52-56, DOI: 10.21437/Interspeech.2016-419. **Nominated for Best Student Paper at Interspeech 2016.**
 - [77] *F. Tao*, L. Daudet, C. Poellabauer, S.L. Schneider, and C. Busso, "A portable automatic PATA-KA syllable detection system to derive biomarkers for neurological disorders," in Interspeech 2016, San Francisco, CA, USA, September 2016, pp. 362-366, DOI: 10.21437/Interspeech.2016-789.
 - [78] *F. Tao*, J.H. L. Hansen, and C. Busso, "Improving boundary estimation in audiovisual speech activity detection using Bayesian information criterion," in Interspeech 2016, San Francisco, CA, USA, September 2016, pp. 2130-2134, DOI: 10.21437/Interspeech.2016-406.
 - [79] *R. Lotfian* and C. Busso, "Retrieving categorical emotions using a probabilistic framework to define preference learning samples," in Interspeech 2016, San Francisco, CA, USA, September 2016, pp. 490-494, DOI: 10.21437/Interspeech.2016-1052.
 - [80] *S. Parthasarathy* and C. Busso, "Defining emotionally salient regions using qualitative agreement method," in Interspeech 2016, San Francisco, CA, USA, September 2016, pp. 3598-3602,

DOI: 10.21437/Interspeech.2016-429.

- [81] *R. Lotfian* and C. Busso, "Practical considerations on the use of preference learning for ranking emotional speech," in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2016), Shanghai, China, March 2016, pp. 5205-5209, DOI: 10.1109/ICASSP.2016.7472670.
- [82] *A. Burmania, M. Abdelwahab,* and C. Busso, "Tradeoff between quality and quantity of emotional annotations to characterize expressive behaviors," in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2016), Shanghai, China, March 2016, pp. 5190-5194, DOI: 10.1109/ICASSP.2016.7472667.
- [83] *A. Jakkam* and C. Busso, "A multimodal analysis of synchrony during dyadic interaction using a metric based on sequential pattern mining," in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2016), Shanghai, China, March 2016, pp. 6085-6089, DOI: 10.1109/ICASSP.2016.7472846.
- [84] T. Hasan, *M. Abdelwahab, S. Parthasarathy,* C. Busso, and Y. Liu, "Automatic composition of broadcast news summaries using rank classifiers trained with acoustic and lexical features," in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2016), Shanghai, China, March 2016, pp. 6080-6084, DOI: 10.1109/ICASSP.2016.7472845.
- [85] *N. Sadoughi* and C. Busso, "Retrieving target gestures toward speech driven animation with meaningful behaviors," in International conference on Multimodal interaction (ICMI 2015), Seattle, WA, USA, November 2015, pp. 115-122, DOI: 10.1145/2818346.2820750.
- [86] A. Iqbal, C. Busso, and N. Gans, "Adjacent vehicle collision warning system using image sensor and inertial measurement unit," in International conference on Multimodal interaction (ICMI 2015), Seattle, WA, USA, November 2015, pp. 291-298, DOI: 10.1145/2818346.2820741.
- [87] *F. Tao, J. Hansen,* and C. Busso, "An unsupervised visual-only voice activity detection approach using temporal orofacial features," in Interspeech 2015, Dresden, Germany, September 2015, pp. 2302-2306, DOI: 10.21437/Interspeech.2015-446.
- [88] *N. Sadoughi,* Y. Liu, and C. Busso, "MSP-AVATAR corpus: Motion capture recordings to study the role of discourse functions in the design of intelligent virtual agents," in 1st International Workshop on Understanding Human Activities through 3D Sensors (UHA3DS 2015), Ljubljana, Slovenia, May 2015, pp. 1-6, DOI: 10.1109/FG.2015.7284885.
- [89] *M. Abdelwahab* and C. Busso, "Supervised domain adaptation for emotion recognition from speech," in International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2015), Brisbane, Australia, April 2015, pp. 5058-5062, DOI: 10.1109/ICASSP.2015.7178934.
- [90] *R. Lotfian* and C. Busso, "Emotion recognition using synthetic speech as neutral reference," in International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2015), Brisbane, Australia, April 2015, pp. 4759-4763, DOI: 10.1109/ICASSP.2015.7178874.
- [91] *M. Abdelwahab* and C. Busso, "Evaluation of syllable rate estimation in expressive speech and its contribution to emotion recognition," in IEEE Spoken Language Technology Workshop (SLT), South Lake Tahoe, CA, USA, December 2014, pp. 472-477, DOI: 10.1109/SLT.2014.7078620.
- [92] *N. Sadoughi,* Y. Liu, and C. Busso, "Speech-driven animation constrained by appropriate discourse functions," in International conference on multimodal interaction (ICMI 2014), Istanbul, Turkey, November 2014, pp. 148-155, DOI: 10.1145/2663204.2663252.
- [93] *N. Li* and C. Busso, "User-independent gaze estimation by exploiting similarity measures in the eye pair appearance eigenspace," in International conference on multimodal interaction (ICMI 2014), Istanbul, Turkey, November 2014, pp. 335-338, DOI: 10.1145/2663204.2663250.
- [94] *F. Tao* and C. Busso, "Lipreading approach for isolated digits recognition under whisper and neutral speech," in Interspeech 2014, Singapore, September 2014, pp. 1154-1158, DOI: 10.21437/Interspeech.2014-294.
- [95] *S. Mariooryad, R. Lotfian,* and C. Busso, "Building a naturalistic emotional speech corpus by retrieving expressive behaviors from existing speech corpora," in Interspeech 2014, Singapore, September 2014, pp. 238-242, DOI: 10.21437/Interspeech.2014-60.
- [96] *N. Li* and C. Busso, "Evaluating the robustness of an appearance-based gaze estimation method for multimodal interfaces," in International conference on multimodal interaction (ICMI 2013), Sydney, Australia, December 2013, pp. 91-98, DOI: 10.1145/2522848.2522876.
- [97] *N. Li* and C. Busso, "Driver mirror-checking action detection using multi-modal signals," in

- The 6th Biennial Workshop on Digital Signal Processing for In-Vehicle Systems, Seoul, Korea, September-October 2013, pp. 101-108.
- [98] *N. Li*, A. Sathyanarayana, C. Busso, and J.H.L. Hansen, "Rear-end collision prevention using mobile devices," in The 6th Biennial Workshop on Digital Signal Processing for In-Vehicle Systems, Seoul, Korea, September-October 2013, pp. 36-43.
 - [99] *S. Mariooryad* and C. Busso, "Analysis and compensation of the reaction lag of evaluators in continuous emotional annotations," in Affective Computing and Intelligent Interaction (ACII 2013), Geneva, Switzerland, September 2013, pp. 85-90, DOI: 10.1109/ACII.2013.21. **Nominated for Best Student Paper at ACII 2013.**
 - [100] *J.P. Arias*, C. Busso, and N.B. Yoma, "Energy and F0 contour modeling with functional data analysis for emotional speech detection," in Interspeech 2013, Lyon, France, August 2013, pp. 2871-2875, DOI: 10.21437/Interspeech.2013-253.
 - [101] *N. Li* and C. Busso, "Analysis of facial features of drivers under cognitive and visual distractions," in IEEE International Conference on Multimedia and Expo (ICME 2013), San Jose, CA, USA, July 2013, pp. 1-6, DOI: 10.1109/ICME.2013.6607575. **Nominated for Best Student Paper at ICME 2013.**
 - [102] *T. Tran*, *S. Mariooryad*, and C. Busso, "Audiovisual corpus to analyze whisper speech," in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2013), Vancouver, BC, Canada, May 2013, pp. 8101-8105, DOI: 10.1109/ICASSP.2013.6639243.
 - [103] *S. Mariooryad* and C. Busso, "Feature and model level compensation of lexical content for facial emotion recognition," in IEEE International Conference on Automatic Face and Gesture Recognition (FG 2013), Shanghai, China, April 2013, pp. 1-6, DOI: 10.1109/FG.2013.6553752.
 - [104] C. Busso and *T. Rahman*, "Unveiling the acoustic properties that describe the valence dimension," in Interspeech 2012, Portland, OR, USA, September 2012, pp. 1179-1182, DOI: 10.21437/Interspeech.2012-124.
 - [105] *S. Mariooryad* and C. Busso, "Factorizing speaker, lexical and emotional variabilities observed in facial expressions," in IEEE International Conference on Image Processing (ICIP 2012), Orlando, FL, USA, September-October 2012, pp. 2605-2608, DOI: 10.1109/ICIP.2012.6467432.
 - [106] D. Tick, *T. Rahman*, C. Busso, and N. Gans, "Indoor robotic terrain classification via angular velocity based hierarchical classifier selection," in IEEE International Conference on Robotics and Automation (ICRA 2012), St. Paul, MN, USA, May 2012, pp. 3594-3600, DOI: 10.1109/ICRA.2012.6225128.
 - [107] *T. Rahman* and C. Busso, "A personalized emotion recognition system using an unsupervised feature adaptation scheme," in International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2012), Kyoto, Japan, March 2012, pp. 5117-5120, DOI: 10.1109/ICASSP.2012.6289072.
 - [108] *J.J. Jain* and C. Busso, "Assessment of driver's distraction using perceptual evaluations, self assessment and multimodal feature analysis," in 5th Biennial Workshop on DSP for In-Vehicle Systems, Kiel, Germany, September 2011.
 - [109] *T. Rahman*, *S. Mariooryad*, *S. Keshavamurthy*, G. Liu, J.H.L. Hansen, and C. Busso, "Detecting sleepiness by fusing classifiers trained with novel acoustic features," in 12th Annual Conference of the International Speech Communication Association (Interspeech-2011), Florence, Italy, August 2011, pp. 3285-3288, DOI: 10.21437/Interspeech.2011-809.
 - [110] X. Fan, C. Busso, and J.H.L. Hansen, "Audio-visual isolated digit recognition for whispered speech," in European Signal Processing Conference (EUSIPCO-2011), Barcelona, Spain, August-September 2011, pp. 1500-1503.
 - [111] *J. Jain* and C. Busso, "Analysis of driver behaviors during common tasks using frontal video camera and CAN-Bus information," IEEE International Conference on Multimedia and Expo (ICME 2011), Barcelona, Spain, July 2011, DOI: 10.1109/ICME.2011.6011950. **Hewlett Packard Best Paper Award at ICME 2011.**
 - [112] C. Busso, A. Metallinou, and S. Narayanan, "Iterative feature normalization for emotional speech detection," in International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2011) Prague, Czech Republic, May 2011, pp. 5692-5695, DOI: 10.1109/ICASSP.2011.5947652.
 - [113] A. Metallinou, C.-C. Lee, C. Busso, S. Carnicke, and S. Narayanan, "The USC CreativeIT database: A multimodal database of theatrical improvisation," in Workshop on Multimodal

- Corpora: Advances in Capturing, Coding and Analyzing Multimodality (MMC 2010), Valletta, Malta, May 2010.
- [114] A. Metallinou, C. Busso, S. Lee, and S. Narayanan, "Visual emotion recognition using compact facial representations and viseme information," in International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2010), Dallas, TX, USA, March 2010, pp. 2474-2477, DOI: 10.1109/ICASSP.2010.5494893.
 - [115] E. Mower, A. Metallinou, C.-C. Lee, A. Kazemzadeh, C. Busso, S. Lee, and S.S. Narayanan, "Interpreting ambiguous emotional expressions," in International Conference on Affective Computing and Intelligent Interaction (ACII 2009), Amsterdam, The Netherlands, September 2009, pp. 1-8, DOI: 10.1109/ACII.2009.5349500.
 - [116] C.-C. Lee, C. Busso, S. Lee, and S.S. Narayanan, "Modeling mutual influence of interlocutor emotion states in dyadic spoken interactions," in Interspeech 2009, Brighton, UK, September 2009, pp. 1983-1986, DOI: 10.21437/Interspeech.2009-480.
 - [117] C.-C. Lee, E. Mower, C. Busso, S. Lee, and S. Narayanan, "Emotion recognition using a hierarchical binary decision tree approach," in Interspeech 2009, Brighton, UK, September 2009, pp. 320-323, DOI: 10.21437/Interspeech.2009-105. **First place, Classifier Sub-Challenge event at the Interspeech 2009 emotion challenge.**
 - [118] C. Busso and S.S. Narayanan, "Scripted dialogs versus improvisation: Lessons learned about emotional elicitation techniques from the IEMOCAP database," in Interspeech 2008 - Eurospeech, Brisbane, Australia, September 2008, pp. 1670-1673, DOI: 10.21437/Interspeech.2008-463.
 - [119] C. Busso and S.S. Narayanan, "The expression and perception of emotions: Comparing assessments of self versus others," in Interspeech 2008 - Eurospeech, Brisbane, Australia, September 2008, pp. 257-260, DOI: 10.21437/Interspeech.2008-92.
 - [120] C. Busso and S.S. Narayanan, "Recording audio-visual emotional databases from actors: a closer look," in Second International Workshop on Emotion: Corpora for Research on Emotion and Affect, International conference on Language Resources and Evaluation (LREC 2008), Marrakech, Morocco, May 2008, pp. 17-22.
 - [121] C. Busso and S.S. Narayanan, "Joint analysis of the emotional fingerprint in the face and speech: A single subject study," in International Workshop on Multimedia Signal Processing (MMSP 2007), Chania, Crete, Greece, October 2007, pp. 43-47, DOI: 10.1109/MMSP.2007.4412814.
 - [122] V. Rozgić, C. Busso, P.G. Georgiou, and S. Narayanan, "Multimodal meeting monitoring: Improvements on speaker tracking and segmentation through a modified mixture particle filter," in International Workshop on Multimedia Signal Processing (MMSP 2007), Chania, Crete, Greece, October 2007, pp. 60-65, DOI: 10.1109/MMSP.2007.4412818.
 - [123] C. Busso, S. Lee, and S.S. Narayanan, "Using neutral speech models for emotional speech analysis," in Interspeech 2007 - Eurospeech, Antwerp, Belgium, August 2007, pp. 2225-2228, DOI: 10.21437/Interspeech.2007-605.
 - [124] C. Busso, P.G. Georgiou, and S. Narayanan, "Realtime monitoring of participants interaction in a meeting using audio-visual sensors," in International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2007), Honolulu, HI, USA, April 2007, vol. 2, pp. 685-688, DOI: 10.1109/ICASSP.2007.366328.
 - [125] C. Busso and S.S. Narayanan, "Interplay between linguistic and affective goals in facial expression during emotional utterances," in 7th International Seminar on Speech Production (ISSP 2006), Ubatuba-SP, Brazil, December 2006, pp. 549-556.
 - [126] M. Bulut, C. Busso, S. Yildirim, A. Kazemzadeh, C.M. Lee, S. Lee, and S. Narayanan, "Investigating the role of phoneme-level modifications in emotional speech resynthesis," in 9th European Conference on Speech Communication and Technology (Interspeech2005 - Eurospeech), Lisbon, Portugal, September 2005, pp. 801-804, DOI: 10.21437/Interspeech.2005-378.
 - [127] C. Busso, S. Hernanz, C.W. Chu, S. Kwon, S. Lee, P.G. Georgiou, I. Cohen, and S. Narayanan, "Smart Room: Participant and speaker localization and identification," in International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2005), Philadelphia, PA, USA, March 2005, vol. 2, pp. 1117-1120, DOI: 10.1109/ICASSP.2005.1415605.
 - [128] N.B. Yoma, C. Busso, J. Inzunza, and F. Huenupan, "Packet-loss modeling with state duration constraints and VoIP based on perceptual quality maximization," in 10th International Confer-

- ence on Speech and Computer (SPECOM 2005), Patras, Greece, October 2005, pp. 757-760.
- [129] C. Busso, Z. Deng, S. Yildirim, M. Bulut, C.M. Lee, A. Kazemzadeh, S. Lee, U. Neumann, and S. Narayanan, "Analysis of emotion recognition using facial expressions, speech and multimodal information," in Sixth International Conference on Multimodal Interfaces ICMI 2004, State College, PA, October 2004, pp. 205-211, ACM Press, DOI: 10.1145/1027933.1027968. **ICMI Ten-Year Technical Impact Award.**
 - [130] Z. Deng, C. Busso, S. Narayanan, and U. Neumann, "Audio-based head motion synthesis for avatar-based telepresence systems," in ACM SIGMM 2004 Workshop on Effective Telepresence (ETP 2004), New York, NY, October 2004, pp. 24-30, ACM Press, DOI: 10.1145/1026776.1026784.
 - [131] C.M. Lee, S. Yildirim, M. Bulut, A. Kazemzadeh, C. Busso, Z. Deng, S. Lee, and S.S. Narayanan, "Emotion recognition based on phoneme classes," in 8th International Conference on Spoken Language Processing (ICSLP 04), Jeju Island, Korea, October 2004, pp. 889-892, DOI: 10.21437/Interspeech.2004-322.
 - [132] S. Yildirim, M. Bulut, C.M. Lee, A. Kazemzadeh, C. Busso, Z. Deng, S. Lee, and S.S. Narayanan, "An acoustic study of emotions expressed in speech," in 8th International Conference on Spoken Language Processing (ICSLP 04), Jeju Island, Korea, October 2004, pp. 2193-2196, DOI: 10.21437/Interspeech.2004-242.
 - [133] N.B. Yoma, J. Hood, and C. Busso, "An UDP-based real time protocol for the internet," in International Telecommunications Symposium (ITS 2002), Natal, Brazil, September 2002.

Book chapters

- [1] C. Pelachaud, C. Busso, and D. Heylen, "Multimodal behavior modeling for socially interactive agents," in The Handbook of Socially Interactive Agents: 20 Years of Research on Intelligent Virtual Agents, Embodied Conversational Agents, and Social Robotics, B. Lugrin, C. Pelachaud, and D. Traum, Eds., pp. 259-310. Association for Computing Machinery, Human-Centered Computing, New York, NY, USA, September 2021, DOI: 10.1145/3477322.3477331.
- [2] S. Jha and C. Busso, "Head pose as an indicator of drivers' visual attention," in Vehicles, Drivers, and Safety, H. Abut, J.H.L. Hansen, G. Schmidt, and K. Takeda, Eds., vol. 2 of Intelligent Vehicles and Transportation, pp. 113-132. De Gruyter, May 2020, DOI: 10.1515/9783110669787-008.
- [3] N. Li and C. Busso, "Driver mirror-checking action detection," in DSP for In-Vehicle Systems and Safety, H. Abut, J.H.L. Hansen, G. Schmidt, K. Takeda, and H. Ko, Eds., Intelligent Vehicles and Transportation, pp. 55-76. DeGruyter, July 2017, DOI: 10.1515/9781501504129-005.
- [4] N. Sadoughi and C. Busso, "Head motion generation," in Handbook of Human Motion, B. Müller, S.I. Wolf, G.-P. Brueggemann, Z. Deng, A. McIntosh, F. Miller, and W. Scott Selbie, Eds., pp. 1-25. Springer International Publishing, January 2017, DOI: 10.1007/978-3-319-30808-1_4-1.
- [5] C.-C. Lee, J. Kim, A. Metallinou, C. Busso, S. Lee, and S.S. Narayanan, "Computational speech processing methods in affective computing: From scientific inquiry to engineering applications," in Handbook of Affective Computing, R. Calvo, S. D'Mello, J. Gratch, and A. Kappas, Eds., pp. 170-183. Oxford University press, New York, NY, USA, December 2014, DOI: 10.1093/oxfordhb/9780199942237.013.021.
- [6] N. Li and C. Busso, "Using perceptual evaluation to quantify cognitive and visual driver distractions," in Smart Mobile In-Vehicle Systems – Next Generation Advancements, G. Schmidt, H. Abut, K. Takeda, and J. H. L. Hansen, Eds., pp. 183-207. Springer, New York, NY, USA, January 2014, DOI: 10.1007/978-1-4614-9120-0_11.
- [7] C. Busso, M. Bulut, and S.S. Narayanan, "Toward effective automatic speech emotion recognition systems," in Social emotions in nature and artifact: emotions in human and human-computer interaction, S. Marsella J. Gratch, Ed. pp. 110-127. Oxford University Press, New York, NY, USA, November 2013, DOI: 10.1093/acprof:oso/9780195387643.003.0008.
- [8] C. Busso and J. Jain, "Advances in multimodal tracking of driver distraction," in DSP for In-Vehicle Systems & Safety, J. Hansen, P. Boyraz, K. Takeda, and H. Abut, Eds., pp. 253-270. Springer, New York, NY, USA, December 2011, DOI: 10.1007/978-1-4419-9607-7_18.
- [9] C. Busso, M. Bulut, S. Lee, and S.S. Narayanan, "Fundamental frequency analysis for speech

- emotion processing,” in *The Role of Prosody in Affective Speech*, Sylvie Hancil, Ed., pp. 309-337. Peter Lang Publishing Group, Berlin, Germany, July 2009.
- [10] C. Busso, Z. Deng, U. Neumann, and S.S. Narayanan, “Learning expressive human-like head motion sequences from speech,” in *Data-Driven 3D Facial Animations*, Z. Deng and U. Neumann, Eds., pp. 113-131. Springer-Verlag London Ltd, Surrey, United Kingdom, 2007, DOI: 10.1007/978-1-84628-907-1_6.

Abstracts

- [1] *K. Rosero, A. N. Salman*, C. Busso, and R. Hallac, “A tailored machine learning approach for cleft lip symmetry analysis,” in *The American Cleft Palate Craniofacial Association (ACPA 2024)*, Denver, CO, April 2024.
- [2] R. Cuenca, M. J. Serafino, G. P. Tortorelli, R. C. Faram, K.E. Higgins, S.S. Khajotia, Y.-S. L. Cheng, J. M. Plemons, V. Woo, C. Busso, *K.R. Caughlin*, and J.A. Jo, “Dual-excitation multispectral autofluorescence lifetime endoscopy for clinical label-free metabolic imaging of oral lesions,” in *Imaging, Therapeutics, and Advanced Technology in Head and Neck Surgery and Otolaryngology 2023*, San Francisco, CA, USA, January-February 2023, vol. SPIE PC12354, pp. 12354-8, DOI: 10.1117/12.2647972.
- [3] E. Duran, S. Cheng, R. Cuenca, B. Ahmed, J. Ji, V.V. Yakovlev, M. Martinez, M. Al-Khalil, H. Al-Enazi, Y.S. L. Cheng, J. Wright, C. Busso, and J. Jo, “Machine-learning driven automated discrimination of precancerous and cancerous from benign oral lesions using multispectral autofluorescence imaging endoscopy,” in *Imaging, Therapeutics, and Advanced Technology in Head and Neck Surgery and Otolaryngology 2022*, San Francisco, CA, USA, March 2022, vol. SPIE PC11935, p. PC1193507, DOI: 10.1117/12.2608843.
- [4] E. Duran, S. Cheng, R. Cuenca, B. Ahmed, J. Ji, V.V. Yakovlev, M. Martinez, M. Al-Khalil, H. Al-Enazi, Y.S. L. Cheng, J. Wright, C. Busso, and J. Jo, “Computer-aided detection system for automated discrimination of precancerous and cancerous from healthy oral tissue based on multispectral autofluorescence lifetime imaging endoscopy,” in *Imaging, Therapeutics, and Advanced Technology in Head and Neck Surgery and Otolaryngology 2022*, San Francisco, CA, USA, March 2022, vol. SPIE PC11935, p. PC1193506, DOI: 10.1117/12.2608818.
- [5] *S. Jha* and C. Busso, “Analysis of head pose as an indicator of drivers’ visual attention,” in *7th Biennial Workshop on DSP for In-Vehicle Systems and Safety*, Berkeley, CA, USA, October 2015,
- [6] C.M. Lee, S. Yildirim, M. Bulut, C. Busso, A. Kazamzadeh, S. Lee, and S. Narayanan, “Effects of emotion on different phoneme classes,” *Journal of the Acoustical Society of America*, vol. 116, pp. 2481, 2004, DOI: 10.1121/1.4784911.
- [7] M. Bulut, S. Yildirim, S. Lee, C.M. Lee, C. Busso, A. Kazamzadeh, and S. Narayanan, “Emotion to emotion speech conversion in phoneme level,” *Journal of the Acoustical Society of America*, vol. 116, pp. 2481, 2004, DOI: 10.1121/1.4784908.
- [8] S. Yildirim, M. Bulut, C. Busso, C.M. Lee, A. Kazamzadeh, S. Lee, and S. Narayanan, “Study of acoustic correlates associate with emotional speech,” *Journal of the Acoustical Society of America*, vol. 116, pp. 2481, 2004, DOI: 10.1121/1.4784909.

ArXiv Papers

- [1] *L. Goncalves, S.-G. Leem, W.-C. Lin*, B. Sisman, and C. Busso, “Versatile audiovisual learning for handling single and multi modalities in emotion regression and classification tasks,” *ArXiv e-prints (arXiv:2305.07216)*, pp. 1-14, May 2023. DOI: 10.48550/arXiv.2305.07216.
- [2] K. Zhou, B. Sisman, C. Busso, and H. Li, “Mixed emotion modelling for emotional voice conversion,” *ArXiv e-prints (arXiv:2210.13756v2)*, pp. 1-5, October 2022, DOI: 10.48550/arXiv.2210.13756.
- [3] *Y. Qiu*, T. Misu, and C. Busso, “Driving anomaly detection using conditional generative adversarial network,” *ArXiv e-prints (arXiv:2203.08289)*, pp. 1-15, March 2022, DOI: 10.48550/arXiv.2203.08289.
- [4] V. Sethu, E. Mower Provost, J. Epps, C. Busso, N. Cummins, and S. Narayanan “The ambiguous world of emotion representation,” *ArXiv e-prints 1909.00360*, pp. 1-19, September 2019,

DOI: 10.48550/arXiv.1909.00360.

- [5] S.-F. Chang, A. Hauptmann, L.-P. Morency, S. Antani, D. Bulterman, C. Busso, J. Chai, J. Hirschberg, R. Jain, K. Mayer-Patel, R. Meth, R. Mooney, K. Nahrstedt, S. Narayanan, P. Natarajan, S. Oviatt, B. Prabhakaran, A. Smeulders, H. Sundaram, Z. Zhang, and M. Zhou, “Report of 2017 NSF workshop on multimedia challenges, opportunities and research roadmaps,” ArXiv e-prints (arXiv:1908.02308), pp. 1-150, August 2019, DOI: 10.48550/arXiv.1908.02308.

INVITED TALKS &
PRESENTATIONS

Busso, Carlos (2023). “Robust Multimodal Machine Learning”, NTHU Internet of Senses Research Center Seminar Series, Seminar at National Tsing Hua University (NTHU), Hsinchu City, Taiwan (Hosted by Prof. Chi-Chun Lee, 12/21/2023).

Busso, Carlos (2023). “Can the Production and Perception of Human Emotions Inspire Speech-Based Affective Computing?”, Keynote at IEEE Automatic Speech Recognition and Understanding Workshop, (Hosted by Prof. Chi-Chun Lee, Prof. Yu Tsao, and Prof. Hsin-Min Wang, 12/19/2023).

Busso, Carlos (2023). “Multimodal machine learning for human-centric tasks”, Center for Language & Speech Processing (CLSP) Seminar, Johns Hopkins University, Baltimore, MD, USA (Hosted by Prof. Najim Dehak, 11/17/2023).

Busso, Carlos (2023). “Advances in multimodal machine learning solutions for speech processing tasks”, IIT Ropar Data Science talk, Virtual seminar (Hosted by Prof. Abhinav Dhall, 6/28/2023).

Busso, Carlos (2023). “Advances in multimodal machine learning solutions for speech processing tasks”, Amazon Alexa Distinguished Speaker Series, Virtual seminar (Hosted by Dr. Yang Liu, 6/29/2023).

Busso, Carlos and Fei Tao (2023). “Advances on Multimodal Machine Learning Solutions for Speech Processing Tasks and Emotion Recognition”, IEEE Signal Processing Society Webinar, (Hosted by Michelle Demydenko and Nanette Januszkiewicz, 01/19/2023).

Busso, Carlos (2022). “Robust Emotion Recognition”, IEEE Signal Processing Society, Bangalore Chapter, at Indian Institute of Science, Bangalore, India (Hosted by Dr. Sriram Ganapathy, 11/09/2022).

Busso, Carlos (2022). “Robust Emotion Recognition”, Invited Tech Talk at Amazon, (Hosted by Dr. Chao Wang, 11/01/2022).

Busso, Carlos (2020). “Multimodal Emotion Recognition: Understanding the production process before modeling multimodal behaviors”, Keynote at the workshop Speech, Music and Mind (SMM2020), (Hosted by Dr. Venkatasubramanian Viraraghavan, 10/23/2020).

Busso, Carlos (2020). “Temporal Head Pose Estimation from Point Cloud in Naturalistic Driving Conditions”, Seminar at Texas Analog Center of Excellence (TxACE) at UT Dallas (Hosted by Prof. Ken O, 10/7/2020)

Busso, Carlos (2020). “Multimodal Assessment of Visual Attention and Driving Anomalies”, Virtual Lunch & Learn seminar at Affectiva (Hosted by Dr. Taniya Mishra, 4/28/2020).

Busso, Carlos (2020). “Multimodal Emotion Analysis and Synthesis: From audiovisual interplay to contextual information in dyadic human interaction”, Keynote at the international conference on Multimodal Modeling (MMM 2020), Deajeon, South Korea (Hosted by Dr. Yong Man Ro, 1/8/2020).

Busso, Carlos (2019). “Driving Anomaly Detection With Conditional Generative Adversarial Network”, Seminar at Texas Analog Center of Excellence at UT Dallas (Hosted by Prof. Ken O, 11/20/2019).

Busso, Carlos, Vidhyasaharan Sethu, Shrikanth Narayanan (2019). “The Ambiguous World of Emotion Representation”, Tutorial at the International Conference on Affective Computing and Intelligent Interaction ACII 2019, Cambridge, UK (9/3/2019).

Busso, Carlos (2019). “Tracking the Behavior and Visual Attention of a Driver Using Multimodal

Sensors in Naturalistic Scenarios”, eSeminar hosted by Texas Analog Center of Excellence (Hosted by Prof. Ken O, 6/21/2019).

Busso, Carlos (2018). “Generating data-driven human-like behaviors for conversational agents”, Keynote at International Workshop on Multimodal Analyses enabling Artificial Agents in Human-Machine Interaction (MA3HMI 2018) (Hosted by Dr. Ronald Böck, 10/16/2018)

Busso, Carlos (2018). “Probabilistic estimation of visual attention”, Seminar at Texas Analog Center of Excellence (TxACE) at UT Dallas (Hosted by Prof. Ken O, 10/10/2018)

Busso, Carlos (2018). “Deep Learning Architectures for Audiovisual Speech Recognition”, Keynote at Multimodal HCI Workshop at Northwestern Polytechnical University (NPU) (Hosted by Prof. Dongmei Jiang, 05/14/2018)

Busso, Carlos (2018). “Novel Formulations and Deep Learning Structures for Robust Speech Emotion Recognition Systems”, Keynote at Apple (Hosted by Dr. Vikramjit Mitra, 05/04/2018)

Busso, Carlos (2018). “Novel Formulations for Speech Emotion Recognition”, Keynote at Cogito Presents: Behavioral Signal Processing and Machine Learning (Hosted by Dr. John Kane, 03/29/2018).

Busso, Carlos (2018). “Tracking Distractions and Visual Attention Using Multiple Noninvasive Sensors”, Seminar at Texas Instrument (Hosted by Dr. Aish Dubey, 02/14/2018).

Busso, Carlos (2017). “Speech emotion recognition: are we there yet?”, Keynote at 2nd International Workshop on Automatic Sentiment Analysis in the Wild (WASA 2017) (10/23/2017).

Busso, Carlos (2017). “Speech emotion recognition: Robustness and generalization”, Seminar at Alibaba (Hosted by Dr. Gang Liu, 07/18/2017).

Busso, Carlos (2016). “Challenges in robust speech emotion recognition in mismatched conditions”, Visit at the Matsuyama lab at Kyoto University, Tokyo, Japon (hosted by Dr. Hiroaki Kawashim and Dr. Rodrigo Verschae, 11/17/2016)

Busso, Carlos (2016). “Robust speech emotion recognition in mismatched conditions”, Statistical Machine Learning Seminar, Richardson, Texas, USA (hosted by Dr. Richard Golden, 04/15/2016).

Busso, Carlos (2016). “Emotion in Group”, Catalyzing Research in Multimodal Learning Analytics Workshop, Bloomington, Indiana (hosted by Dr. Marcelo Worsley and Dr. Cindy Hmelo-Silver, 02/21/2016).

Busso, Carlos (2016). “Multimodal Signal Processing: Current Research Directions”, UTSW/UTD Medical Simulation and Training Workshop, Richardson, Texas, USA (hosted by Dr. Ann Majewicz, 01/22/2016).

Busso, Carlos (2015). “Feature normalization and model adaptation for robust speech emotion recognition in mismatched conditions”, Seminar at The University of New South Wales, Advanced Signal Processing Workshop, Sydney, Australia (hosted by Dr. Julien Epps, 04/17/2015).

Busso, Carlos (2014). “Multimodal Analysis, Recognition and Synthesis of Expressive Human Behaviors”, Seminar at Koç Üniversitesi, Istanbul, Turkey (hosted by Dr. Engin Erzin, 11/11/2014).

Carlos Busso (2014). “Tracking Driver Distractions Using Multiple Noninvasive Sensors”, Seminar at Philips, Eindhoven, The Netherlands (hosted by Dr. Murtaza Bulut, 11/10/2014).

Provost, Emily Mower and Carlos Busso (2014). “Computational Models for Audiovisual Emotion Perception”, TUTORIAL at INTERSPEECH 2014, Singapore (09/14/2014).

Busso, Carlos (2014). “Using Neutral Models and Contextual Information to Improve Speech Emotion Recognition Systems”, Professorial talk at Columbia University, New York, NY 10027 (hosted by Dr. Julia Hirschberg, 07/02/2014).

Busso, Carlos (2014). “Multimodal Analysis, Recognition and Synthesis of Expressive Human Behaviors”, VASC Seminar at Carnegie Mellon University, Pittsburgh, PA 15213 (hosted by Dr. Kris Kitani, 06/30/2014).

- Busso, Carlos (2014). "Compensation of Lexical and Speaker variability for Emotional Recognition", Seminar at The University of Pennsylvania at Philadelphia, PA 19104, USA (hosted by Dr. Ani Nenkova, 01/27/2014).
- Busso, Carlos (2013). "Improving the Robustness of Emotional Speech Detection Systems", Seminar at The University of Texas A&M, College Station, TX 77843, USA (hosted by Dr. Ricardo Gutierrez-Osuna, 11/18/2013).
- Busso, Carlos (2013). "Improving the Robustness of Emotional Speech Detection Systems", Seminar at The University of Texas at San Antonio, San Antonio, TX 78249, USA (hosted by Dr. Ram Krishnan, 11/01/2013).
- Busso, Carlos (2013). "Modeling of Driver Behavior in Real World Scenarios Using Multiple Non-invasive Sensors", AI Seminar at The University of Michigan, Ann Arbor, MI 48109, USA (hosted by Dr. Emily Mower Provost, 10/22/2013).
- Busso, Carlos (2013). "Keep Your Eyes on the Road, Your Hands Upon the Wheel", Tech on Tap Seminars, Trinity Hall Irish Pub & Restaurant, Dallas TX 75206, USA (hosted by Beth Keithly, 03/06/2013).
- Busso, Carlos (2012). "Improving the Robustness of Emotional Speech Detection Systems", Seminar at The University of Texas at Arlington, Arlington TX 76019, USA (hosted by Dr. Ioannis Schizas, 04/06/2012).
- Busso, Carlos (2011). "Tracking expressive nonverbal behavior: a multimodal approach", FLASH (Friday Seminars in Speech, Language and Hearing) Brown Bag Series at the Callier Center, The University of Texas at Dallas, Dallas TX 75235, USA (hosted by Dr. William F. Katz, 10/21/2011).
- Busso, Carlos (2011). "Recognition and synthesis of human behaviors: A multimodal approach", Seminar at Texas Instrument, Dallas, TX, USA (hosted by Dr. Branislav Kisanin, 04/08/2011).
- Busso, Carlos (2011). "Recognition and synthesis of human behaviors: A multimodal approach", Dallas Chapter of IEEE Signal Processing Society Seminar, The University of Texas at Dallas, Richardson, TX, USA (hosted by Dr. Nasser Kehtarnavaz, 02/23/2011).
- Busso, Carlos, John H.L. Hansen (2010). "Advances in Human Assessment: (i) Tracking Nonverbal Behavior (ii) Speaker Variability for Speaker ID", 2010 Fall Seminar Series, The Center for Language and Speech Processing (CLSP) at Johns Hopkins University, Baltimore, MD, USA (hosted by Dr. Hynek Hermansky, 10/26/2010).
- Busso, Carlos (2009). "New Directions in the Automatic Recognition of Emotion in Speech". Seminar at The University of Texas at El Paso, El Paso, TX, USA (hosted by Dr. David Novick and Nigel Ward, 10/21/2009).
- Busso, Carlos (2009). "New directions on automatic emotion speech recognition", Seminar at The University of North Texas, Denton, TX, USA New directions on automatic emotion speech recognition (hosted by Dr. Oscar Garcia, 10/15/2009).
- Busso, Carlos (2009). "Multimodal Analysis of Expressive Human Communication: Speech and gesture interplay", Seminar at The University of Texas at Dallas, Richardson, TX, USA (hosted by Dr. John Hansen and Dr. Hlaing Minn, 03/24/2009).
- Busso, Carlos (2009). "Multimodal Processing of Human Behavior in Intelligent Instrumented Spaces: A Focus on Expressive Human Communication", Information Science Institute (ISI) Natural Language Seminar talks, Marina del Rey, CA, USA (hosted by Dr. Sujith Ravi, 02/27/2009).
- Busso, Carlos (2009). "Multimodal Processing of Human Behavior in Intelligent Instrumented Spaces: A Focus on Expressive Human Communication" Seminar at Nokia, Santa Monica, CA, USA (hosted by Dr. Lance Williams, 01/08/2009).
- Busso, Carlos (2008). "Multimodal Processing of Human Behavior in Intelligent Instrumented Spaces: A Focus on Expressive Human Communication" Microsoft, Redmond, WA, USA. (hosted by Dr. Zhengyou Zhang, 05/05/2008).

LICENSE REVENUE: Note: \$348,979.00 in total license revenue.

MSP-Podcast Corpus	\$314,979.00
MSP-Conversation Corpus	\$18,000.00
MSP-IMPROV Corpus	\$16,000.00

GRANTS & CONTRACTS: Note: \$5,568,404.64 in total research funding, \$4.1M as PI. The list only includes UTD's portion for collaborative projects.

Active Research Projects:

2024-2024 Laboratory for Analytic Sciences (NC State University); \$110,198
"Language-specific adaptation for cross-lingual speech emotion recognition"
PI: Carlos Busso
January 1, 2024 - December 31, 2024

2023-2028 UTSW Interagency Cooperation Contract (ICC); \$174,066.00
"Face the Future: Assessing Speech Articulation in Pediatric Craniofacial Patients through Machine Learning"
PI: Carlos Busso
September 1, 2023 - August 31, 2028

2020-2024 National Science Foundation (NSF) - CNS: 2016719; \$1,075,386
"CCRI: New: Creating the largest speech emotional database by leveraging existing naturalistic recordings"
PI: Carlos Busso
REU supplement, 2021-2022 (\$16,000)
September 1, 2020 - August 31, 2024

2019-2024 National Institute of Health (NIH)- R01:1R01MH122367-01 ; \$189,813
"SCH: INT: Collaborative Research: Passive Sensing of Social Isolation and Loneliness: a Digital Phenotyping Approach"
co-PI: Carlos Busso (UTD sub-contract, in collaboration with Dr. Daniel Fulford (PI) at BU, David Gard at UCSC, Jukka-Pekka Onnela at Harvard)
September 23, 2019 - August 31, 2024

2018-2024 National Institute of Health (NIH)- R01:5R01CA218739-04 ; \$180,363
"Endogenous Fluorescence Lifetime Endoscopy for Early Detection of Oral Cancer/dysplasia"
co-PI: Carlos Busso (UTD sub-contract, in collaboration with Dr. Javier Jo at OU)
April 1, 2018 - August 31, 2024

Completed Research Projects:

2023-2023 Laboratory for Analytic Sciences (NC State University); \$103,628
"Improving Speech Emotion Recognition with Task-Agnostic and Task-Depend Self-Supervised Learning"
PI: Carlos Busso
January 1, 2023 - December 31, 2023

- 2022-2022 Laboratory for Analytic Sciences (NC State University); \$85,430
 “Ordinal speech emotion recognition: Rank-order emotional speech with preference learning ”
 PI: Carlos Busso
 January 1, 2022 - December 31, 2022
- 2020-2023 National Science Foundation (NSF) - IIP: 1950249; \$80,000
 “STTR Phase I: A Self-Learning Approach for In-Vehicle Driver and Passenger Monitoring Through a Sensor Fusion Approach”
 PI: Rajesh Narasimha, Senior personal: Naofal Al-Dhahir and Carlos Busso
 March 3, 2020 - November 30, 2020
- 2019-2020 Honda Research Institute (HRI) ; \$97,227
 “Driving anomaly detection based on multi-modal fusion of driver behavioral signals and external scene information”
 PI: Carlos Busso
 February 1, 2019 - May 21, 2021
- 2020-2021 NEC Foundation Research Grant ; \$50,000
 “Multimodal engagement analysis of students for online education”
 PI: Carlos Busso
 February 15, 2020 - February 14, 2021
- 2018-2019 Honda Research Institute (HRI) ; \$86,079.64
 “Driver anomaly detection based on driver/vehicle sensor signals”
 PI: Carlos Busso
 September 17, 2018 - September 13, 2019
- 2018-2019 NEC Foundation Research Grant ; \$50,000
 “Tracking facial emotional variations on videos: separating lexical and emotional information with novel deep learning solutions”
 PI: Carlos Busso
 September 1, 2018 - August 30, 2019
- 2018-2020 National Science Foundation (NSF) - CNS: 1823166; \$99,390
 “CRI: CI-P: Creating the Largest Speech Emotional Database by Leveraging Existing Naturalistic Recordings”
 PI: Carlos Busso
 REU supplement, 2019-2020 (\$8,000)
 REU supplement, 2018-2019 (\$8,000)
 September 1, 2018 - February 28, 2021
- 2018-2020 SRC/Texas Analog Center of Excellence: Task 2810.014 ; \$180,000
 “Deep Learning Solutions for ADAS: From Algorithms to Real-World Driving Evaluations”
 PI: Carlos Busso, co-PI: Naofal Al-dhahir
 January 1, 2018 - December 31, 2020

- 2017-2020 National Science Foundation (NSF) - IIS: 1718944; \$494,116
 “RI: Small: Integrative, Semantic-Aware, Speech-Driven Models for Believable Conversational Agents with Meaningful Behaviors”
 PI: Carlos Busso
 REU supplement, 2020-2021 (\$8,000)
 REU supplement, 2019-2020 (\$8,000)
 REU supplement, 2018-2019 (\$8,000)
 REU supplement, 2017-2018 (\$8,000)
 September 1, 2017 - August 31, 2021
- 2017-2018 SRI / The Combating Terrorism Technical Support Office (CTTSO) ; \$239,951
 “Intrinsic Voice Variability Factors and Speaker Recognition: Detection and Assessment”
 PI: John H.L. Hansen , Co-PI: Carlos Busso
 January 2, 2018 - March 30, 2019
- 2017-2018 Biometric Center of Excellence (BCOE) ; \$290,189
 “Automatic Audio Stream Processing to Address Diverse Mismatch Scenarios for Speaker Verification”
 PI: John H.L. Hansen , Co-PI: Carlos Busso
 June 1, 2017 - November 30, 2018
- 2016-2017 Biometric Center of Excellence (BCOE) ; \$202,264
 “Speaker Variability - Advancements in Detection and Knowledge Integration of Emotion, Task Stress, Vocal Effort for Speaker Verification in Naturalistic Environments”
 PI: John H.L. Hansen , Co-PI: Carlos Busso
 March 1, 2016 - November 31, 2017
- 2015-2020 National Science Foundation (NSF) - IIS: 1453781; \$495,853
 “CAREER: Advanced Knowledge Extraction of Affective Behaviors During Natural Human Interaction”
 PI: Carlos Busso
 REU supplement, 2019-2020 (\$8,000)
 REU supplement, 2018-2019 (\$8,000)
 REU supplement, 2017-2018 (\$8,000)
 REU supplement, 2016-2017 (\$8,000)
 REU supplement, 2015-2016 (\$8,000)
 September 1, 2015 - August 31, 2021
- 2016-2017 Microsoft Research; \$30,000
 “Support for Data Collection”
 PI: Carlos Busso
 March 1, 2016 - February 28, 2017
- 2015-2016 National Science Foundation (NSF) - IIS: 1540944; \$11,040
 “FG 2015 Doctoral Consortium: Travel Support for Graduate Students”
 PI: Carlos Busso
 April 1, 2015 - March 31, 2016
- 2014-2016 Robert Bosch LLC; \$30,000
 “Detecting Emotionally Salient Speech Segments for Speech Summarization”
 PI: Carlos Busso
 September 1, 2014 - May 30, 2015

- 2014-2017 National Science Foundation (NSF) - IIS: 1450349; \$19,732
 “EAGER: Feasibility of Using Speech as Biomarker for Concussions”
 PI: Carlos Busso (UTD sub-contract, in collaboration with Dr. Christian Poellabauer, PI at University of Notre Dame)
 September 1, 2014 - August 31, 2017
- 2013-2014 Samsung Telecommunications America, LLC; \$120,146
 “SRA: Farsi Speech Recognition Project”
 co-PI: Carlos Busso (PI: John Hansen)
 January 1, 2014 - December 31, 2014
- 2013-2016 National Science Foundation (NSF) - IIS: 1352950; \$82,552
 “EAGER: Investigating the Role of Discourse Context in Speech-Driven Facial Animations”
 PI: Yang Liu, co-PI: Carlos Busso
 September 1, 2013 - February 28, 2016
- 2013-2016 National Science Foundation (NSF) - IIS: 1346655; \$17,840
 “WORKSHOP: Doctoral Consortium for the International Conference on Multimodal Interaction (ICMI 2013)”
 PI: Carlos Busso
 July 1, 2013 - June 30, 2016
- 2013-2014 National Science Foundation (NSF) - IIS: 1329659; \$59,338
 “EAGER: Exploring the Use of Synthetic Speech as Reference Model to Detect Salient Emotional Segments in Speech”
 PI: Carlos Busso
 March 15, 2013 - August 31, 2014
- 2012-2014 Samsung Telecommunications America, LLC; \$441,898
 “Advancements in Automatic Speech Recognition: Corpus Development, Model Training, Dialect/Accent, and Hands-Free Interaction”
 co-PI: Carlos Busso (PI: John Hansen)
 November 1, 2012 - December 31, 2014
- 2012-2016 National Science Foundation (NSF) - IIS: 1217104; \$201,573
 “RI: Small: Collaborative Research: Exploring Audiovisual Emotion Perception using Data-Driven Computational Modeling”
 PI: Carlos Busso (Emily Mower Provost, University of Michigan, Ann Arbor, MI)
 September 1, 2012 - August 31, 2016
- 2012-2014 National Science Foundation (NSF) - IIS: 1249319; \$14,587
 “WORKSHOP: Doctoral Consortium for 14th International Conference on Multimodal Interaction”
 PI: Carlos Busso
 August 1, 2012 - July 31, 2014
- 2011-2012 Samsung Telecommunications America, LLC; \$151,745
 “Standardization of Advanced User Interface for Mobile Devices”
 PI: Carlos Busso
 September 15, 2011 - September 14, 2012

GRADUATE
ADVISEE:

Current Students:

PhD Students Ali N. Salman, “Face emotion recognition”, August 2018 - present
Kayla Caughlin, “Oral cancer detection”, August 2020 - present
Lucas Goncalves, “Audiovisual emotion recognition”, August 2020 - present
Seong-Gyun Leem, “Speech processing in noisy environment”, August 2020 - present
Abinay Reddy Naini, “Speech emotion recognition,” January 2022 -present
Luz Martinez-Lucas, “Speech emotion recognition”, May 2022 - present
Pravin Mote, “Multimodal emotion recognition”, August 2022 - present
Karen Rosero, “Face analysis for healthcare applications”, January 2023 - present

PhD Student Alumni:

- [1] Andrea Vidal (2018-2023) – Openstream.AI, Bridgewater, NJ
PhD Thesis: “Synthesizing lip movements and modeling human behavior by using generative models and graph neural networks”
- [2] Wei-Cheng (Winston) Lin (2019-2023) – Robert Bosch, Pittsburgh, PA
PhD Thesis: “Improving Temporal Modeling and Generalization for Speech Emotion Recognition”
- [3] Yuning Qiu (2018-2022) – Intel, Austin, TX
PhD Thesis: “Unsupervised Driving Anomaly Detection in Naturalistic Driving Scenarios”
- [4] Kusha Sridhar (2017-2021) – Rice University, Houston, TX
PhD Thesis: “Unsupervised Personalization and Deep Uncertainty Modeling for Speech Emotion Recognition”
- [5] Sumit Jha (2016-2021) – Samsung Research America, Irvine, CA
PhD Thesis: “Novel Modeling of Driver Attention in Real World Scenarios using Probabilistic Salient Maps”
- [6] Srinivas Parthasarathy (2015-2019) – Amazon, Sunnyvale, CA
PhD Thesis: “Novel Frameworks for Attribute-Based Speech Emotion Recognition using Time-Continuous Traces and Sentence-Level Annotations”
- [7] Mohammed AbdelWahab (2013-2019) –AT&T Labs Research, Bedminster, NJ
PhD Thesis: “Domain Adaptation for Speech Based Emotion Recognition.”
- [8] Fei Tao (2013-2018) – Uber, San Francisco, CA
PhD Thesis: “Advances in audiovisual speech processing for robust voice activity detection and automatic speech recognition”
- [9] Reza Lotfian (2013-2018) – Cogito, Boston, MA
PhD Thesis: “Machine Learning Solutions for Emotional Speech: Exploiting the Information of Individual annotations.”
- [10] Najmeh Sadoughi (2013-2017) – EMR.AI, San Francisco, CA
PhD Thesis: “Synthesizing Naturalistic and Meaningful Speech-Driven Behaviors.”
- [11] Nanxiang (Sean) Li (2011-2015) – Honda Research Institute, Mountain View, CA
PhD Thesis: “Modeling of Driver Behavior in Real World Scenarios using Multiple Noninvasive Sensors.”
- [12] Soroosh Mariooryad (2010-2014) – Google, Mountain View, CA
PhD Thesis: “Improving Robustness of Emotion Recognition Systems: Continuous Emotion Descriptors, Contextual Information and Lexical Compensation.”

Visiting Scholars and Post-Doctoral Researcher Alumni:

- [1] Huang-Cheng (David) Chou (2021-2022) – National Tsing Hua University
- [2] Yakup Kutlu (2013) – Mustafa Kemal University, Hatay, Turkey
- [3] Youngkwon Lim (2011-2012) – Samsung Telecommunications America, Richardson, TX
- [4] Juan Pablo Arias (2010-2011) –Adexus, Santiago, Chile

Ms. Student Alumni:

- [1] Lucas Goncalves (2020-2023) – PhD student at UT Dallas, Richardson, TX
- [2] Alec Burmania (2016-2018) – Lennox International, , TX
Ms. Thesis: “Methods and experimental design for collecting emotional labels using crowdsourcing.”
- [3] Yunjie Zhang (2015-2017) – PhD student at UT Dallas, Richardson, TX
- [4] Anil Jakkam (2015-2016) – Knowles Intelligent Audio, Mountain View, CA
Ms. Thesis: “A Multimodal Analysis of Synchrony during Dyadic Interaction Using a Metric Based on Sequential Pattern Mining.”
- [5] Sumit Jha (2015-2016) – PhD student at UT Dallas, Richardson, TX
Ms. Thesis: “Analysis and Estimation of Driver Visual Attention using Head Position and Orientation in Naturalistic driving Conditions.”
- [6] Srinivas Parthasarathy (2012-2014) – PhD student at UT Dallas Richardson, TX
Ms. Thesis: “Relation Between Emotion Classification Performance and Evaluator Agreement with Absolute and Relative Descriptors.”
- [7] Sheldon Dsouza (2012) – Apple, Cupertino, CA
- [8] Tauhidur Rahman (2010-2012) – PhD student at Cornell University, Ithaca, NY
Ms. Thesis: “Improving robustness of emotional speech detection system.”
- [9] Shalini Keshavamurthy (2010-2011) – UtopiaCompression Corp, Los Angeles, CA
- [10] Jinesh Jain (2010-2011) – Ford Motors Company, Palo Alto, CA
Ms. Thesis: “Driver Distraction: Multimodal analysis and modeling of driver behavior in real-world scenarios.”
- [11] Anand Batlagundu (2010) – Qualcomm, San Diego, CA
- [12] Somu Palaniappan (2010) – NSN - Nokia Solutions and Networks, Irving Texas
- [13] Premkumar Sridhar (2009) – Ericsson Inc, Dallas, TX

Undergraduate Alumni:

- [1] Shruthi Subramaniam (2020-2023)
- [2] Isaac Brooks (2020-2023)
- [3] Olivia Bonin (2021)
- [4] Luz Martinez-Lucas (2018-2021)
- [5] Jarrod Luckenbaugh (2019-2021)
- [6] Kayla Caughlin (2019-2020)
- [7] Tiancheng Hu (2018-2020)
- [8] John Harvill (2018-2019)
- [9] Elizabeth Higgins (2018-2019)
- [10] Asim Gazi (2017-2018)
- [11] Dorothy Mantle (2017-2018)
- [12] Michelle Bancroft (2016-2018)
- [13] Alec Burmania (2012-2016)
- [14] Jaejin Cho (2014 - 2015)
- [15] Preston Luthy (2015)
- [16] Zackary R Lindstrom (2013- 2014)
- [17] Tam Tran (2011 - 2013)

CLASSROOM TEACHING:	2024	Spring	EE3302.001	Signals and Systems
	2023	Fall	ENGR2300.501	Linear Algebra for Engineers
	2023	Spring	EE3302.001	Signals and Systems
	2022	Fall	ENGR2300.501	Linear Algebra for Engineers
	2022	Spring	EESC6360.001	Digital Signal Processing I

2021	Fall	ENGR2300.004	Linear Algebra for Engineers
2021	Spring	EESC6360.001	Digital Signal Processing I
2020	Fall	EE3302.003	Signals and Systems
2020	Spring	ENGR2300.HON	Linear Algebra for Engineers
2020	Spring	EESC6360.001	Digital Signal Processing I
2019	Fall	EE3302.002	Signals and Systems
2019	Spring	ENGR2300.HON	Linear Algebra for Engineers
2019	Spring	EESC6360.001	Digital Signal Processing I
2018	Fall	EE3302.002	Signals and Systems
2018	Spring	ENGR2300.502	Linear Algebra for Engineers
2018	Spring	EESC6368.001	Multimodal Signal Processing
2017	Fall	ENGR2300.006	Linear Algebra for Engineers
2017	Spring	ENGR2300.002	Linear Algebra for Engineers
2017	Spring	ENGR2300.502	Linear Algebra for Engineers
2016	Fall	EESC6360.502	Digital Signal Processing I
2016	Spring	ENGR2300.002	Linear Algebra for Engineers
2016	Spring	EESC6360.501	Digital Signal Processing I
2015	Fall	EESC6360.502	Digital Signal Processing I
2015	Summer	ENGR2300.0U1	Linear Algebra for Engineers
2015	Spring	EESC6368.001	Multimodal Signal Processing
2015	Spring	ENGR2300.002	Linear Algebra for Engineers
2014	Fall	EESC6360.001	Digital Signal Processing I
2014	Spring	ENGR2300.001	Linear Algebra for Engineers
2014	Spring	ENGR2300.002	Linear Algebra for Engineers
2013	Fall	EESC6349.001	Random Processes
2013	Spring	ENGR2300.001	Linear Algebra for Engineers
2013	Spring	ENGR2300.002	Linear Algebra for Engineers
2012	Fall	EESC6349.001	Random Processes
2012	Summer	ENGR2300.0U1	Linear Algebra for Engineers
2012	Spring	ENGR2300.002	Linear Algebra for Engineers
2012	Spring	EE7V85.501	Special Topics in Signal Proc. - Multimodal Signal Processing
2011	Fall	ENGR2300.003	Linear Algebra for Engineers
2011	Spring	ENGR2300.002	Linear Algebra for Engineers
2010	Fall	ENGR2300.001	Linear Algebra for Engineers
2010	Spring	EE7V85.501	Special Topics in Signal Proc.- Multimodal Signal Processing