Indian Institute of Science and
The IEEE Signal Processing Society, Bangalore Chapter

Cordially invites you to the following talk on

"Robust Emotion Recognition"

Speaker: Prof. Carlos Busso, University of Texas, Dallas (UTD)
Date and Time: 9th November 2022 at 2:00pm**, Refreshments:** 3:00pm
Venue: MMCR (Room No. C241, 1st Floor, Dept. of Electrical Engineering

Abstract of the talk

It is challenging to achieve robust and well-generalized models for tasks involving subjective concepts such as emotion. This tech talk will describe novel approaches to effectively develop robust *speech emotion recognition* (SER) systems. At the resource level, we will describe our effort to collect the MSP-Podcast corpus, which is a large, naturalistic emotional database. The data collection protocol combines machine-learning algorithms to retrieve recordings conveying balanced emotional content annotated with a cost-effective crowdsourcing protocol. To improve the temporal modeling of SER systems, this seminar will also discuss a novel dynamic chunking approach that maps the sequences of different lengths into a fixed number of chunks that have the same duration by adjusting their overlap. This simple chunking procedure creates a flexible framework, facilitating parallel computing. The approach can incorporate different feature extractions and sentence-level temporal aggregation approaches to cope, in a principled way, with a sequence-to-one SER task. Likewise, the seminar will discuss multimodal pre-text tasks that are carefully designed to learn better representations for predicting emotional cues from speech, leveraging the relationship between acoustic and facial features. Finally, the seminar will discuss our current effort to design multimodal emotion recognition strategies that effectively combine auxiliary networks, a transformer architecture, and an optimized training mechanism for aligning modalities, capturing temporal information, and handling missing features. These models offer principled solutions to increase the generalization and robustness of emotion recognitions systems.

Biography

Carlos Busso received his PhD degree (2008) in electrical engineering from the University of Southern California (USC), Los Angeles, in 2008. He is a professor at the Electrical Engineering Department of The University of Texas at Dallas (UTD). At UTD, he leads the Multimodal Signal Processing (MSP) laboratory [http://msp.utdallas.edu]. He is a recipient of an NSF CAREER Award. In 2014, he received the ICMI Ten-Year Technical Impact Award. In 2015, his student received the third prize IEEE ITSS Best Dissertation Award (N. Li). He also received the Hewlett Packard Best Paper Award at the IEEE ICME 2011 (with J. Jain), and the Best Paper Award at the AAAC ACII 2017 (with Yannakakis and Cowie). He received the Best of IEEE Transactions on Affective Computing Paper Collection in 2021 (with R. Lotfian) and in 2022 (with Yannakakis and Cowie). He is the co-author of the winner paper of the Classifier Sub-Challenge event at the Interspeech 2009 emotion challenge. His research interest is in human-centered multimodal machine intelligence and applications. His current research includes the broad areas of affective computing, multimodal human-machine interfaces, nonverbal behaviors for conversational agents, in-vehicle active safety system, and machine learning methods for multimodal processing. His work has direct implication in many practical domains, including national security, health care, entertainment, transportation systems, and education. He was the general chair of ACII 2017 and ICMI 2021. He is a member of ISCA, AAAC, and a senior member of ACM and IEEE.