Abstract:

Emotions have a huge contribution to the development of human beings and decoding them from brain signals has a wide range of practical implications. Decoding emotions is challenging as emotions are complex, vary with time, and involve cognitive processing. In this work, we investigate wavelet, functional connectivity, and graph-theory based features for emotion decoding as these features reflect energy characteristics of an individual electrode, interaction between a pair of electrodes, and structural properties of the brain networks, respectively. We use the DEAP dataset, which comprises of 32-channel EEG data from 32 participants. The data were recorded when they watched music videos of one-minute duration, and each participant self-reported felt emotions on arousal, valence, dominance, liking, and familiarity scales. We perform binary classification of high/low categories of arousal, valence, and dominance. We observe that wavelet features perform better for arousal, connectivity features for valence, and connectivity and graph-theory based features for dominance. This shows that a particular type of feature is not sufficient to adequately and discriminatively represent different emotional states. Therefore, we propose to fuse these features and results show that the feature-level fusion of these features performs better than individual features. With the proposed feature-level fusion, a relative improvement of 3.39% for arousal, 1.48% for valence, and 1.81% for dominance are achieved as compared to considered individual features. And a relative improvement of 6.16% for arousal and 18.98% for valence is observed as compared to the traditional PSD feature originally used for the DEAP dataset.

Next, to induce emotions, music is extensively used as stimuli, and it is well-known that parameters, such as liking and familiarity influence induced emotions. However, their effect on emotion decoding performance has not been studied. Thus, we investigate the emotion decoding performance in different categories of these parameters. The proposed fusion significantly improves the performance as compared to individual features for all cases of familiarity and liking. The relative improvement in performance is approximately 1-5% for familiarity and liking and 1-8% for familiarity and liking combined. We found that decoding of arousal and dominance is improved by this, but we do not find such improvement for valence. The results reveal the importance of considering these parameters while selecting stimuli to induce emotions. Next, we ask if there is any coherent emotional neural representation among the categories of liking and familiarity. The cross-classification framework is employed where EEG data of one category is used for training and other categories for testing. We found that neural-representations are coherent for liking and familiar parameters for arousal.

Lastly, we investigate the time-course of induced emotional states by taking into consideration the level of familiarity and liking. We find that the performance linearly increases with time for arousal for familiar and unfamiliar–high-liking songs; valence for unfamiliar and unfamiliar–high-liking songs; dominance for high-liking songs. And performance linearly decreases with time for valence for unfamiliar–low-liking songs. These trends are statistically significant ($p<0.05$). This analysis provides a novel insight into the underlying mechanisms of induced emotions. Altogether, this work demonstrates the decoding and understanding of music video induced emotional states.