Joyful or Nervous? A Dataset of Awkward, Embarrassed and Uncomfortable Smiles

Justin Heer, Jie Yan, Angelica Lim
Simon Fraser University

Abstract

In the realm of human-agent interaction, smiles are often considered positive social signals; however, smiles represent an entire spectrum of emotion including nervousness, embarrassment and discomfort. The subtle nuances separating joyful smiles from the latter often go undetected, resulting in them being classified as the same. Further development in discerning the differences between smiles would aid researchers in building an emotionally receptive and intelligent virtual agent, able to connect with humans at an emotional level. While datasets containing Duchenne [2] and posed smiles exist[1], there are currently no video datasets specifically for negative-affect smiles. In this paper we describe the detectable social signals of negative smiles and the data collection procedures followed during the creation of an in-the-wild dataset containing 111 short video clips of negative-affect smiles along with an equal number of positive-affect smiles, with high inter-rater reliability (Cohen's kappa=0.81). This publicly available dataset includes full YouTube video sources which can improve context-based detection methods; while other datasets focus on several emotions, this video dataset provides a granular social-signal level assessment of negative-affect smiles. We utilized OpenFace, a deep learning facial action unit(AU) detection algorithm, to extract features from training examples and analyze negative-affect smile composition. We found statistically significant differences in AU12 (Lip Corner Puller, $p = 0.0011$) and AU26 (Jaw Drop, $p < 0.001$) occurrence, as well as noteworthy discrepancies between AU6, AU7 and AU17 when examining negative smiles. Based on research pertaining to discerning between real and false smiles [3], we built a pre-trained CNN + Bi-LSTM classifier, and compared its performance with an SVM classifier using Dynamic Time Warping distance. The tuned SVM-DTW model obtained an F1-score of 67% for negative smile detection, a baseline for machine level accuracy in regards to our dataset, compared to an F1-score of 58% with the CNN-BiLSTM. Finally, we provide suggestions for areas of future research and development, these include utilizing an IRNN, facial alignment, scene context, further data collection as well as the possibility of undertaking a multi-modal approach by investigating the impact of auditory features.

Keywords— Affective Computing, Computer Vision, smiles, Duchenne, false smiles, nervous smiles, negative smiles

References

[1] Hamdi Dibeklioglu, Albert Ali Salah, and Theo Gevers. "Recognition of Genuine Smiles". In: *IEEE Transactions on Multimedia* 17.3 (2015), pp. 279–294. DOI: 10.1109/tmm.2015.2394777.

[2] Paul Ekman, Richard J. Davidson, and Wallace V. Friesen. "The Duchenne smile: Emotional expression and brain physiology: II." In: *Journal of Personality and Social Psychology* 58.2 (1990), pp. 342–353. DOI: 10.1037/0022-3514.58.2.342.

[3] Minh Tu Nguyen et al. "Towards recognizing facial expressions at deeper level: Discriminating genuine and fake smiles from a sequence of images". In: *2019 6th NAFOSTED Conference on Information and Computer Science (NICS)* (2019). DOI: 10.1109/nics48868.2019.9023790.

word count: 334