Recent Debates Regarding Facial Expression Technologies

- Affectiva, November 2022

Over the course of the last year or so, there has been a thread of debate in the media regarding the validity and ethics of facial emotion recognition. This has often reflected the point of view of some data privacy groups who are concerned about the use of facial technologies across several use cases, or the opinions of commercial interests who offer alternative biometric technologies, or traditional research methodologies.

Scrutiny of emerging technologies is vital, and the concerns raised are important points for debate. Affectiva has led the development of the Emotion AI field for over a decade, and the use of automated facial expression analysis in particular. We have taken great pains during that time to ensure the technology we offer is both ethical, accurate and meaningful. The very extensive use of our technology across the industry (70% of the world's largest advertisers) is testament to the success of that effort. As a result of this work, these criticisms are not relevant to the technology used by Affectiva, and as used by the market research industry, particularly for the evaluation of audience response to content, advertising, and brand experiences. This document explains why.

Recent criticisms have centred on 3 areas:

1. Face "reading" technology is unethical and an invasion of privacy
2. Face "reading" technology is inaccurate and potentially biased against different groups
3. Inferring emotional states from facial expressions is not based on sound science, as expressions do not have a universal meaning or one-to-one relationship with emotional states, and their meaning varies significantly by context. (This mainly draws on work done by Lisa Feldman-Barrett and her collaborators [1])

These criticisms are important considerations, but are not valid with regards to Affectiva's Media analytics work for the following reasons, and we deal with each in turn:

1. Face reading technology is unethical and an invasion of privacy

    - Our data collection is done with explicit opt-in consent by research participants, every time. This is a condition of use of our technology. Affectiva does not analyse data collected without consent.
    - Affectiva's emotion AI does not and cannot recognise people, only expressions – a person's identity is noise we need to ignore to recognise facial expressions across a diverse range of people and populations. This is not a Facial

Identification technology; it is a facial expression technology, which is materially different in nature.

- Our technology does not infer demographic information from the face data – we do not estimate gender, age, ethnicity and so on. Where this data is included in our system it is self-reported by research participants.

2. Face reading technology is inaccurate and potentially biased against different groups

- Affectiva's technology has been upgraded many times over the last 10 years and each time we take care to train our algorithms on highly diverse datasets. We have access to more real-world face data than any other organisation – over 13 billion frames of data from over 90 countries, and as a result we train on predominantly non-Caucasian datasets with excellent balance in gender and age. We take great care to ensure equal accuracy across demographic groups and so can be confident that our system exhibits minimal bias.
- We accurately replicate the perceptions of facial movements that people recognise – specifically using the Facial Action Coding System (FACS) [2, 3] – which is accepted as the gold standard method of identifying facial movements in the academic community [4]. The robustness of the Affectiva method has been validated and published in peer reviewed and open-source articles [5, 6].
- We train our algorithms separately for different use cases – which is critical for ensuring accuracy in specific contexts – for instance we use different algorithms for driver detection in the automotive sector vs our market research applications using webcam and smartphones. This is not a generic system trying to work across many uses – which is often the case with free-to-use technology that some users may have tried.
- Affectiva and many of our clients have independently validated the technology and demonstrated relationships between people's facial expressions when they watch content, and their subsequent behaviour – e.g. facial expression responses to advertising has a relationship with subsequent behaviours and the sales effectiveness of that advertising among the broad population [7, 8, 9, 10, 11]. This would not happen if the technology was inaccurate or meaningless.
- Affectiva's technology is used by over 600 academic institutions, who still find it useful and powerful as a measurement of people's reactions – again suggesting that the idea that there is no scientific consensus about the utility of facial data is incorrect.

3. Inferring emotional states from facial expressions is not based on sound science, as expressions do not have a universal meaning or one-to-one relationship with emotional states, and their meaning varies significantly by context.

    - We agree that context is very important–a smile can mean happiness in one context but embarrassment in another. This is why our metrics are trained specifically for different contexts (see above), and critically, our data is interpreted with the knowledge of the stimulus that the person is referring to. Affectiva's metrics are interpreted by analysts rather than simply yielding a score which is blindly followed, and we are in full agreement that a system that is blind to context may yield misleading results. Put simply – when using our system, the analyst knows if the research participant is watching something funny, embarrassing, or likely to cause offence, and can judge the meaning of the facial reactions observed appropriately.
    - The Affectiva system primarily focuses on detecting specific facial movements, such as smiles, furrowed brows, raised eyebrows and so on, based on the Facial Action Coding System (FACS). We identify 20 different Action Units (AUs) – these facial movements are the building blocks for a wide range of facial expressions and are universal across different people. The universal presence of these action units across cultures has not been challenged. Critically, within our system, the meaning of these facial actions is for interpretation by the analyst.
    - Affectiva's metrics are also interpreted in the context of the country of data collection – our system includes normative comparisons for over 90 markets worldwide, so analytics can allow for the fact that some cultures are more expressive on different dimensions than others.
    - In many cases, analysts also have access to what people say about an ad or experience, as well as the facial data. Our results dashboards allow analysts to look at facial expressions based on whether people said they liked an ad or film, for instance – which allows for a rounded view.
    - While recent studies have highlighted that there is more cultural nuance in facial expression than the original science in this space suggested [2, 12], this does not mean that there is no universality in expressions. The academic debate centres on the extent to which there is variation across context, rather than suggesting that there is no universality to expressions. The Ekman FACS system, which our algorithms replicate, is still in widespread use in academia worldwide, and the base science behind classification of facial movements that is still robust. In addition, more recent work shows that when context is held constant, the meaning of different expressions remains consistent across cultures (e.g. Cowen et al in their recent Nature paper [13], showed that a

wide variety of expressions remain consistent in similar contexts worldwide). And finally, our own analysis of the huge number of facial videos we have processed confirms that the same sets of expressions appear repeatedly across cultures and markets in response to content. The assertion that there is no cultural consistency to facial expressions is not supported by the evidence.

- We would agree that given that there is contextual variation in the meaning of specific expressions, it is reasonable to ask whether it is right to label certain combinations of expressions with emotion labels, as the academic EMFACS system does. As a result, our primary metrics focus on specific expressions (e/g smile, brow furrows, downturned mouths), which analysts interpret based on context. However, in response to demands from the academic community and users, we do also include combinations of these expressions which are labelled in line with EMFACS classifications in our tools. We remain of the view that these can be useful generalisations and indicate a reasonable hypothesis about an individual's response. However, we are clear that in some circumstances the reason for a particular expression may differ to the EMFACS label. Our guidance to users makes clarifies this, and as ever we encourage users to interrogate the data appropriately to confirm if the generalisation makes sense in that specific context.
- We are concerned that there are double standards being applied in this debate. While it is possible that facial signals may vary depending on context, this is also true of almost all biometric and survey methods. The reasons for a Galvanic Skin Response spike, or EEG signal change, or Heart-rate variation, or a survey answer, can be myriad, and there is no single meaning to any of those signals. While some reviewers may challenge the consistency of meaning of facial signals, there are also reviewers who challenge the consistency of other biometric signals, e.g. [14]. Indeed, there also papers that find facial expressions to be more useful measures of emotion than other biometric signals, e.g. [15]. So the importance of context applies to almost all consumer data, not just facial signals, and the utility of other biometric signals has also been placed in doubt.
- Finally, the assertion that there is no meaningful information to be gleaned from people's faces about their reactions is entirely at odds with everyday human experience, where we look to non-verbal cues such as expressions all the time to help us understand each other. Technologies such as Affectiva's emotion AI are simply allowing that insight to be gleaned in a scalable way. Facial signals also have the advantage over other biometric methods that they do not require a trained neuroscientist to interpret, with little chance that a research user will be able to sense-check those answers. In contrast, people have an intuitive understanding of the meaning of a wide range facial signals as they

use them every day and can evaluate their meaning far more readily as a result.

Summary

Understanding the meaning of facial signals is an evolving field and we encourage industry and academia to include signals from multiple modalities for the most accurate picture. But we continue to find that automated Facial expression analysis, if conducted with consent, is both ethical, accurate, and can meaningfully tell us something about human responses. Contextual awareness is necessary to avoid simplistic conclusions, but as a tool to understand people, unobtrusively and at scale, this technology has enormous power and value.

References

[1] Barrett, Lisa Feldman, et al. "Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements." Psychological science in the public interest 20.1 (2019): 1-68.

[2] Ekman, Paul, and Wallace V. Friesen. "Facial action coding system." Environmental Psychology & Nonverbal Behavior (1978).

[3] Rosenberg, Erika L., and Paul Ekman, eds. What the face reveals: Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS). Oxford University Press, 2020.

[4] Martinez, Brais, et al. "Automatic analysis of facial actions: A survey." IEEE transactions on affective computing 10.3 (2017): 325-347.

[5] McDuff, Daniel, et al. "AFFDEX SDK: a cross-platform real-time multi-face expression recognition toolkit." Proceedings of the 2016 CHI conference extended abstracts on human factors in computing systems. 2016.

[6] Bishay, Mina, et al. "AFFDEX 2.0: A Real-Time Facial Expression Analysis Toolkit." arXiv preprint arXiv:2202.12059 (2022). Accepted at the FG2023 conference.

[7] McDuff, Daniel, et al. "Predicting ad liking and purchase intent: Large-scale analysis of facial responses to ads." IEEE Transactions on Affective Computing 6.3 (2014): 223-235.

[8] Koldra, Evan, et al. Do emotions in advertising drive sales? https://ana.esomar.org/documents/do-emotions-in-advertising-drive-sales--8059.

[9] McDuff, Daniel, and Rana El Kaliouby. "Applications of automated facial coding in media measurement." IEEE transactions on affective computing 8.2 (2016): 148-160.

[10] Teixeira, Thales, Rosalind Picard, and Rana El Kaliouby. "Why, when, and how much to entertain consumers in advertisements? A web-based facial tracking field study." Marketing Science 33.6 (2014): 809-827.

[11] McDuff, Daniel, et al. "Automatic measurement of ad preferences from facial responses gathered over the internet." Image and Vision Computing 32.10 (2014): 630-640.

[12] Ekman, Paul, and Wallace V. Friesen. "Constants across cultures in the face and emotion." Journal of personality and social psychology 17.2 (1971): 124.

[13] Cowen, Alan S., et al. "Sixteen facial expressions occur in similar contexts worldwide." Nature 589.7841 (2021): 251-257.

[14] Aleksandra Kołodziej, Mikołaj Magnuski, Anastasia Ruban, Aneta Brzezicka “No relationship between frontal alpha asymmetry and depressive disorders in a multiverse analysis of five studies” eLife 10:e60595 (2021)

[15] Mohammad Soleymani et al “Analysis of EEG Signals and Facial Expressions for Continuous Emotion Detection” IEEE Transactions on Affective Computing, VOL. 7, NO. 1 (2018)