

Machine versus Human Affect Recognition of Posed and Spontaneous Expressions

Hyunwoo Kim, Jeffrey Girard, Dennis Küster, & Eva Krumhuber

University College London
University of Kansas
University of Bremen



01 INTRODUCTION

The field of automated facial expression analysis (AFEA) has progressed rapidly in recent years. Existing research suggests that machine classifiers perform reasonably well in recognising emotions from posed and prototypical expressions (Zeng, Pantic, Roisman & Huang, 2009). However, much less is known about the reliability of AFEA in the context of subtle and spontaneous expressions (Yitzhak et al., 2018). Furthermore, there is limited evidence as to whether peak intensity images extracted from dynamic expressions (e.g., Lewinsky et al., 2014) appropriately reflect the target emotion. The present study aims to fill this gap in the literature by investigating machine versus human affect recognition for peak intensity frames of posed and spontaneous expressions.

Research Questions:

1. Does the machine classifier outperform humans in recognising affect from peak intensity images?
2. Are Posed expressions more recognisable than spontaneous expressions?

02 METHOD

Videos of posed ($N=108$) and spontaneous ($N=54$) facial expressions showing the six basic emotions (i.e., anger, disgust, fear, happiness, sadness, and surprise; Ekman, 1999) were taken from Krumhuber et al. (2020). The FACET software (iMotions, 2016) was used to identify the image frame from each video that had the highest evidence value for the target emotion (i.e., the emotion that the participant was instructed to pose or reported feeling). These images were extracted programmatically (Python 3) and shown to new study participants (i.e., raters). 154 white/Caucasian participants (76 females, $M_{age} = 29.43$) from the UK rated the images on the extent (0% - not at all, 100% - extremely) to which the six basic emotions are recognisably expressed in the image. To compare machine evidence values with human confidence ratings, FACET's evidence scores were converted into probabilities by using the formula provided in the FACET documentation ($1/(1 + (10^{\wedge}\text{evidence}))$; iMotions, 2016), and then into confidence odds ($1/((1/p)-1)$).

04 DISCUSSION

- ✓ The machine achieved recognition rates similar (and often superior) to those of human observers
- ✓ Consistent with previous findings, posed expressions were better recognised than spontaneous expressions
- ✓ The peak intensity frames extracted by the machine were sufficient to detect emotions in the face
- ✓ Future research should compare human vs. machine performance in both static and dynamic facial expressions

03 RESULTS

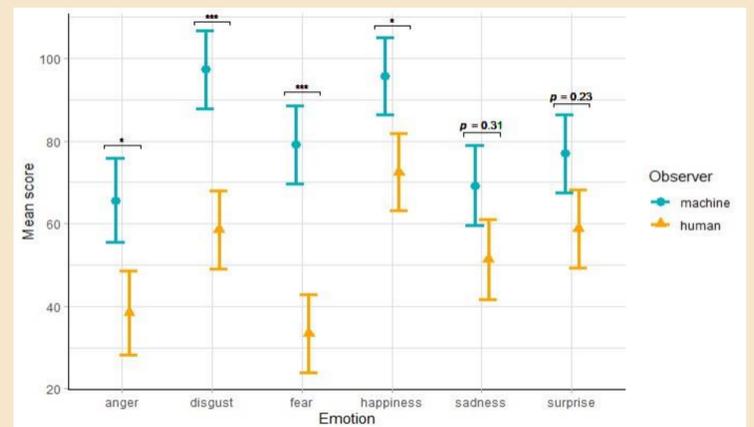


Figure 1. Machine versus Human Affect Recognition of the Six Basic Emotions

Recognition scores for the peak intensity images were submitted to a 2 (rater: machine vs. human) \times 6 (emotion: anger, disgust, fear, happiness, sadness, surprise) ANOVA which revealed a significant interaction between the two factors, $F(5, 312) = 2.82$, $p = 0.016$, $\eta_p^2 = .04$. Tukey's HSD post hoc tests showed that recognition scores by the machine were significantly higher for anger ($MD = 27.2$, $p = 0.012$), disgust ($MD = 38.8$, $p < .001$), fear ($MD = 45.8$, $p < .001$) and happiness ($MD = 23.2$, $p = 0.031$) compared to human observers (see Figure 1).

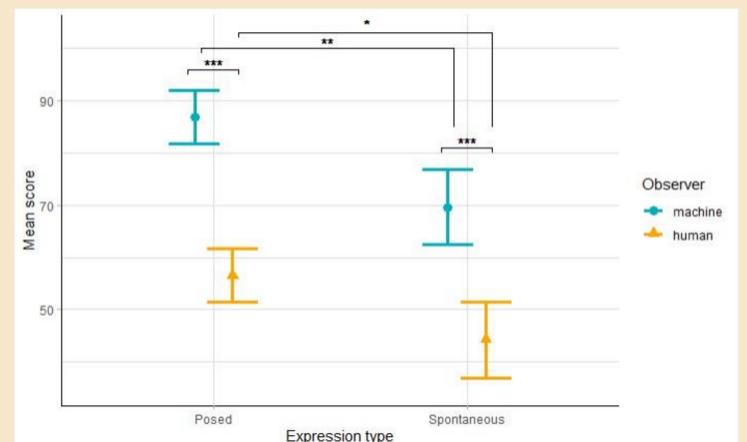


Figure 2. Machine versus Humans Affect Recognition of Posed and Spontaneous Expressions

To compare recognition rates between posed and spontaneous expressions, a 2 (rater: machine vs. human) \times 2 (expression type: posed vs. spontaneous) ANOVA was carried out. A significant main effect of classifier $F(1, 320) = 90.93$, $p < .001$, $\eta_p^2 = .22$, revealed that the machine ($M = 81.17$, $SD = 31.09$) outperformed humans ($M = 52.49$, $SD = 24.31$) in emotion recognition. Furthermore, there was a significant main effect of expression type $F(1, 320) = 21.57$, $p < .001$, $\eta_p^2 = .06$, in the sense that posed expressions were better recognised ($M = 71.77$, $SD = 28.62$) than spontaneous expressions ($M = 56.95$, $SD = 34.24$). No interaction effect was observed between classifier and type of expressions (see Figure 2).

References:

- Ekman, P. (1999). Basic emotions. *Handbook of cognition and emotion*, 98(45-60), 16.
- iMotions (2016). *Biometric Research Platform 5.7, Emotion FACET*, iMotions A/S, Copenhagen, Denmark.
- Krumhuber, E. G., Küster, D., Namba, S., & Skora, L. (2020). Human and machine validation of 14 databases of dynamic facial expressions. *Behavior research methods*, 1-16.
- Lewinsky, P., den Uyl, T. M., & Butler, C. (2014). Automated facial coding: Validation of basic emotions and FACS AUs in FaceReader. *Journal of Neuroscience, Psychology, and Economics*, 7(4), 227.
- Yitzhak, N., Giladi, N., Gurevich, T., Messinger, D. S., Prince, E. B., Martin, K., & Aviezer, H. (2017). Gently does it: Humans outperform a software classifier in recognizing subtle, nonstereotypical facial expressions. *Emotion*, 17(8), 1187.
- Zeng, Z., Pantic, M., & Huang, T. S. (2009). Emotion recognition based on multimodal information. In *Affective information processing* (pp. 241-265). Springer, London.