(Mis)decoding affect from faces and brains. Some cautionary notes

Marco Viola (Roma Tre University)

Despite technical progress, automatic systems aimed at "decoding" a subject's affective states based on objective measures, such as patterns of facial movements or neural activity, are undermined by intricate epistemological and theoretical issues. A common problem is that they often inherit the following tenets of (one version of) Paul Ekman's Basic Emotion Theory: (i) the ontology of emotions for all mankind comprises only six categories (anger, disgust, fear, happiness, sadness, surprise); (ii) instances of emotion belonging to each category can be distinguished based on distinctive markers, such as a dedicated facial expression and neural marker.
In my talk, after having succinctly sketched Ekman's theoretical commitments, I will show that, despite being increasingly criticized on theoretical grounds (often with sound reasons), it still looms at large in practice, i.e. in psychological and neuroscientific research, as well as in state-of-art automated system for decoding affect.
Then, I will briefly examine two such systems, namely facial emotion recognition (FER) and functional magnetic resonance imaging (fMRI), examining some of their shortcomings, with a particular emphasis on the shortcomings stemming from their reliance on the abovementioned tenets (i) and (ii).
I conclude by considering and comparing the ethical risks posed by the two techniques.