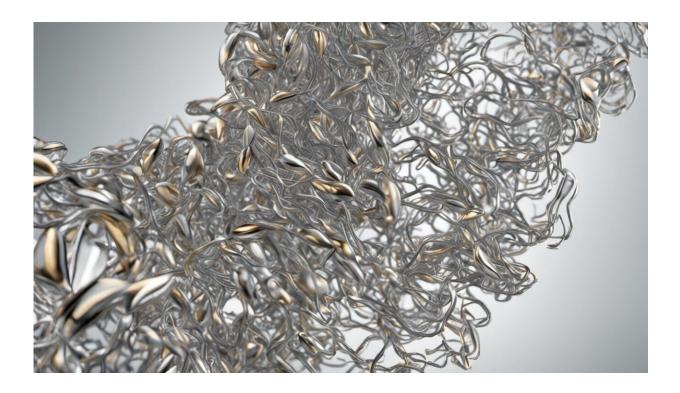# Can computers be trained to understand body language?

September 5 2019, by Jessica Hallman

Credit: AI-generated image (disclaimer)

Humans are able to "read" others' body language for cues on their emotional state. For instance, noticing that a friend is nervous by their tapping foot, or that a loved one who is standing tall feels confident. Now, a team of researchers at Penn State are exploring if computers can be trained to do the same.

The team is investigating whether modern computer vision techniques could match the cognitive ability of humans in recognizing bodily expressions in real-world, unconstrained situations. If so, these capabilities might allow for a large number of innovative applications in areas including [information management](#) and retrieval, [public safety](#), [patient care](#) and social media, the researchers said.

"Computers and robots in the future will be interacting with more people," said James Wang, professor in the College of Information Sciences and Technology (IST) and a member of the research team. "Today's computers, to a large extent, merely follow orders. In the future, robots and computers will act more like partners to humans and work together. And to do so, they'll need to understand their emotions."

College of IST doctoral candidate Yu Luo, working with Wang and other faculty on the team, processed a large number of movie clips and built a dataset of more than 13,000 human characters with nearly 10,000 body movements. According to the researchers, studies have shown that the human body may be more diagnostic than the face in recognizing [human emotion](#).

"The term in psychology is called 'socio-editing,'" said Luo. "People can use it to manipulate their facial expression, but it's much more difficult to control their body. The body language projects different emotions."

Next, the researchers used computer vision methods to locate and track each person across different frames in the scene, ultimately marking each individual in a clip with a unique ID number. Finally, the researchers utilized crowdsourced human annotators to review the movie clips and identify the emotion of each individual featured in one of 26 categorical emotions, i.e., peace, affection, esteem, anticipation, engagement, confidence, happiness, pleasure, excitement, surprise, sympathy, confusion, disconnection, fatigue, embarrassment, yearning,

disapproval, aversion, annoyance, anger, sensitivity, sadness, disquietment, fear, pain and suffering, as well as in three dimensions of emotion, i.e., valence, arousal and dominance.

"We found that interpreting emotion based on body language is complex," said Wang. "There are a lot of subtleties that we are trying to understand. Even for humans there are a lot of inconsistencies.

"People don't agree with each other when it comes to interpreting emotions," he added. "You may think a person is happy, I may think they're excited, and perhaps both of us are correct. There's often no ground truth, which makes data-driven modeling highly challenging."

Once the researchers built the dataset and applied the human-perceived [emotional](#) annotations for each individual, they used state-of-the-art statistical techniques to validate their quality-control mechanisms and thoroughly analyzed the consensus level of their verified data labels. Further, they constructed automated emotion-recognition systems from human skeletons and image sequences. Specifically, deep learning techniques and hand-crafted, Laban-movement analysis-based features demonstrated effectiveness for the task.

They found that the [computer](#) model could identify arousal, or how energized the experience feels, with a high level of precision. However, the researchers also found that humans are better than computers at identifying the valence —how negative or positive the experience feels.

The current results were made possible by a seed grant from the College of IST and ongoing research is supported by a recent award from the Amazon Research Award Program. The team also recently was awarded a planning project from the National Science Foundation to build a community to develop the data infrastructure to be utilized in this research.

Wang and Luo worked with other Penn State researchers on the project, including Jianbo Ye, former doctoral student and lab mate in the College of IST; Reginald Adams and Michelle Newman, professors of psychology; and Jia Li, professor of statistics. A provisional patent application recently has been filed, and the work will be published in a forthcoming issue of the *International Journal of Computer Vision*.

"The barrier of entry for this line of research is pretty high," said Wang. "You have to use knowledge from psychology, you have to develop and integrate data science methods, and you have to use statistical modeling to properly collect affective data. This shows that we are at the forefront of sciences and technology in this important information subdomain."

More information: Yu Luo et al. ARBEE: Towards Automated Recognition of Bodily Expression of Emotion in the Wild, *International Journal of Computer Vision* (2019). DOI: 10.1007/s11263-019-01215-y

Provided by Pennsylvania State University