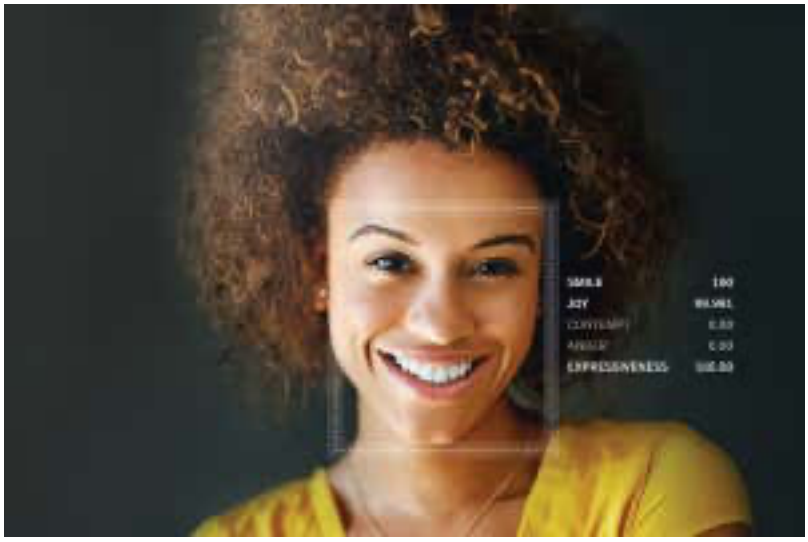
Emotion-detecting AIs are here – do they work and how should we feel?

Several companies have built AIs designed to figure our when you are angry or sad or excited. But there are serious questions about their accuracy, and the extent to which they should be used in public life



Some AIs attempt to identify people's emotions from their facial expressions

RANA EL KALIOUBY was alone in her flat, messaging her husband. "How are you doing?" he typed. "I'm fine," she typed back. Except that wasn't true. The couple had been apart for weeks and she was feeling miserable. Had he been in the room, he could have read the emotions on her face at a glance. But he was miles away.

It is a scene that could easily have played out during a coronavirus pandemic lockdown, when colleagues, friends and even families were cut off from one another. But it actually took place 20 years ago, soon after el Kaliouby had moved from Egypt to the UK to study, leaving her husband behind.

It was in that moment, she says, that she realised how technology was blind to human emotions. Ever since, el Kaliouby has dreamed of building an emotionally intelligent computer – or as she puts it "a mind-reading machine". With so many relationships mediated by text or video call these days, it is a technology that couldn't be more relevant.

These days, the company el Kaliouby co-founded, Affectiva, and others like it, claim to have systems capable of detecting human emotions. The promises they make about the potential of this emotion artificial intelligence (AI) are staggering. Computers, they say, will know if we are distracted while driving, angrily typing an email that we may regret or when our mental health is beginning to slump. In fact, systems like this already exist. But do they live up to their billing? And do we really want machines that know how we feel?

To make her dream a reality, el Kaliouby is relying on AI systems that learn from data how to do certain tasks. Take an AI that recognises cat pictures. It is trained by being shown lots of photos of cats and, in the process, learns what shapes, colours and other features to look out for. To begin with, the system is told whether the pictures contain cats or not. Eventually, given enough examples, it can learn to discern this itself.

Like spotting a cat, recognising a human emotion is a skill that, in principle, could be mastered by AI. Far more is at stake, though, at least for el Kaliouby. "These technologies that interface with us on a day-to-day basis need to know human, they need to understand human," she says. We all know that video calls and messaging apps can't fully replicate the experience of a face-to-face chat and part of that is because it is harder to remotely read other people's emotions. Yet computer-mediated conversations are set to become more common, especially as more of us work from home. One possible application of emotion AI is to help us judge whether an audience is bored, excited or somewhere in between.

The field of affective computing – making machines that can recognise, interpret and simulate human emotions – has been going for years yet is in some ways still in its infancy. We are a long way from machines that can realistically simulate our emotions. But when it comes to machines that can recognise our feelings, we appear to be getting there.

Since Affectiva was founded in 2009, it has mostly worked with marketing firms to help them understand how audiences react to videos and other marketing content. This is often done by recording videos of people's faces as they encounter a poster, say, or watch an advert. By tracking how facial expressions change when they encounter these things, Affectiva says it can predict how successful an ad campaign will be. It isn't alone. Among others, Realeyes, a company headquartered in London, has also worked with advertisers to measure how attentive viewers are.

Other companies reckon they can glean insights from the sound of our voices alone. Behavioral Signals, a tech firm in Los Angeles, has developed software to classify emotion based on a speaker's tone. Its algorithm tracks pitch, volume, rhythm, intonation and other features of speech. "We focus on how something is being said," says chief executive Rana Gujral. "Oftentimes we don't even convert the audio into text."

This kind of intelligence could be handy for businesses, not least firms that field lots of phone calls. Behavioral Signals claims to be able to tell within 30 seconds of a call whether a customer will take a particular action, such as commit to paying off a debt.

The company says it has sold its tech to a major European bank, which used tone-of-voice analysis to match callers with staff in call centres. Irate customers were automatically triaged to especially calm and collected handlers. There was a 20 per cent improvement in call outcomes, says Gujral, which translated into an expected $300 million, had the bank relied on the system across its business for a whole year.

Hidden signals

If a car could recognise its passengers' emotions, it might make driving safer

But Behavioral Signals is going even further. Gujral says he is working with another company, which he won't name, that hopes to use AI in order to pick up vocal signals associated with depressive symptoms. The idea is to predict the likelihood of someone going on to attempt suicide based on their tone of voice during conversations with a psychologist or caregiver. This is experimental, but "they're actually working on live patients right now", says Gujral.

Clearly, emotional AI could be useful. Can algorithms really grasp human emotions, though? It isn't easy to evaluate the various companies' claims directly because their algorithms typically aren't made public. Even so, the science of emotion recognition can offer some insights.

We know it is possible to classify facial expressions. In the mid-20th century, psychologist Paul Ekman pinpointed facial movements that he argued could be associated with broad emotional states. There are now considered to be seven such emotional states under Ekman's Facial Action Coding System (FACS): anger, fear, disgust, happiness, sadness, surprise and contempt. These are often referred to as the "universal emotions" and well-practised humans can get good at identifying them from facial signals alone, with one study suggesting they get it right 77 per cent of the time.

But there is a rather large catch: accurately identifying the "anger" state like this doesn't necessarily mean the person displaying it is angry. One researcher who doubts the usefulness of this approach is Lisa Feldman Barrett, a psychologist at Northeastern University in Massachusetts. She and her colleagues recently completed a large review of existing research to find out whether there was any relationship between specific facial expressions and internal emotional states. The gist of their findings is that the evidence is scant – a scowl can be associated with a wide range of emotions, from anger to confusion to concentration. What's more, the use of various expressions varies noticeably between cultures. We all like to think we could be as perceptive as Lady Macbeth: "Your face, my thane, is as a book where men may read strange matters", but fully decoding expressions isn't so easy in reality.

If humans aren't perfect, can AI do better? Well, FACS is still an important ingredient in the algorithms designed by AI companies, including Affectiva and Realeyes. This makes Barrett highly sceptical of them. She argues that emotions are more like "episodes": there is no one brain signal for "anger" and there isn't one recognisable way of expressing that emotion on the face either.

Similar quibbles plague voice-based emotion recognition. Margaret Lech at RMIT University in Melbourne, Australia, and her colleagues reviewed a series of studies on the accuracy of such systems and found that they were able to correctly identify emotions around 60 per cent of the time on average – hardly a glowing result.

Some proponents of emotion AI accept that this is valid argument and admit that the systems only recognise expressions or intonations, not necessarily the true emotions beneath. Rosalind Picard, who co-founded Affectiva and is now director of the Affective Computing group at the Massachusetts Institute of Technology, is one of them. "I am not proposing one could measure affective state directly, but rather measure observable functions of such states," she wrote in a 1995 paper.

Others see the distinction as a straw man. Barrett's critique might hold if a person or AI observes a face with no contextual information, they would argue, but that isn't what happens in real life. Emotion AI is typically used to see how a person reacts to a specific thing, like a funny video clip. In that context, a grin is probably just a grin: an uncomplicated indication of amusement.

"Emotion AI could tell us whether an audience is bored, excited or somewhere in between"

Wild feelings

What happens as this technology begins to be used more widely, though? Some companies working in affective computing have made their tech available to study and this has led to the discovery of some examples of bias. Lauren Rhue at the University of Maryland catalogued the results of emotion AIs developed by Microsoft and Chinese tech firm Megvii when the systems were fed pictures of white and black basketball players. "Both services interpret black players as having more negative emotions than white players," she wrote in a 2018 paper. New Scientist asked both firms for a response to these claims. Megvii says that it puts great importance on fairness and doesn't use its algorithms in scenarios that don't meet its standards. Microsoft declined to comment.

Such ethical considerations aren't stopping some police forces from trialling emotion AI in the wild. Lincolnshire Police in the UK recently hit the news because it has received government funding to deploy a system intended to detect the emotions of people captured on CCTV footage. This will, for now, be limited to a trial, and footage will be deleted after 31 days, according to the force.

Researchers at New York University's AI Now Institute recently questioned how fair it is to use emotion-detecting AI on members of the public without explicit consent. Among the systems mentioned in the institute's 2019 annual report is one offered by Oxygen

Forensics, a US-based, Russian-owned firm that sells software to the FBI, London's Metropolitan Police and Interpol. Among its products, the company offers facial recognition technology that it claims can detect emotion. The institute's report argued that there was little to no scientific basis for this technology in general and said its use in areas including criminal justice was "concerning". Lee Reiber at Oxygen Forensics has previously said that being able to detect anger, stress or anxiety can help investigators.

"A scowl can be associated with a range of emotions, from anger to concentration to confusion"

Laughter or despair?

"I don't think we should be using emotion AI broadly in public life, particularly to make decisions that have legal consequence for people," says Jevan Hutson at the University of Washington School of Law in Seattle. Hutson likens emotion AI to phrenology, the discredited 19th-century idea that bumps on people's heads were correlated with personality traits. He says countries should adopt policies and laws to restrict the use of emotion AI in certain areas. Law enforcement, job recruitment, surveillance in public places – all of these applications and more could be deeply problematic if we get this stuff wrong, he says.

Meanwhile, there are plenty of people working to soothe such issues by making emotion AI more accurate. The most obvious way to do this is to provide the AI with more context. Imagine an AI trying to work out whether a grimace signifies laughter or despair. If it had the tools to recognise whether a person is hanging out with friends or strangers – perhaps by recognising the other faces nearby – then it could more confidently settle on laughter. This isn't perfect, of course, but in general the more contextual information an AI has, the more likely it is to draw accurate inferences.

Take this to its logical conclusion and we might end up with emotion AI systems that hoover up information on our voices, our body movements, faces and data about our environment, who we hang out with and what the people around us are doing.

"I don't think we should be using emotion AI in public life, particularly in decisions of legal consequence"

This might sound a long way off, but there is one place where we are already seeing the first steps in this direction: cars. El Kaliouby and her colleagues at Affectiva are among the engineers now developing AI to monitor the behaviour of vehicle occupants. "We've expanded to things like activity detection," she says. "Are you holding a cellphone to your ears? Are the kids fighting in the back seat?"

Ultimately, such a system could form the basis of an autonomous car that even knows when to take control of the wheel. Vehicle maker Toyota is already building a prototype car equipped with an emotion AI system developed by US-based SRI International. Other car-makers are working with Affectiva, el Kaliouby says, and she expects to launch the first commercial versions of the technology in two to three years.

All this leaves computers that can detect our emotions at a pivotal juncture. They still aren't entirely trusted, and yet they are on the cusp of becoming far more widespread. Society hasn't fully grappled with the consequences. But one thing is for sure: we have come a long way since el Kaliouby sent those lonely messages two decades ago.