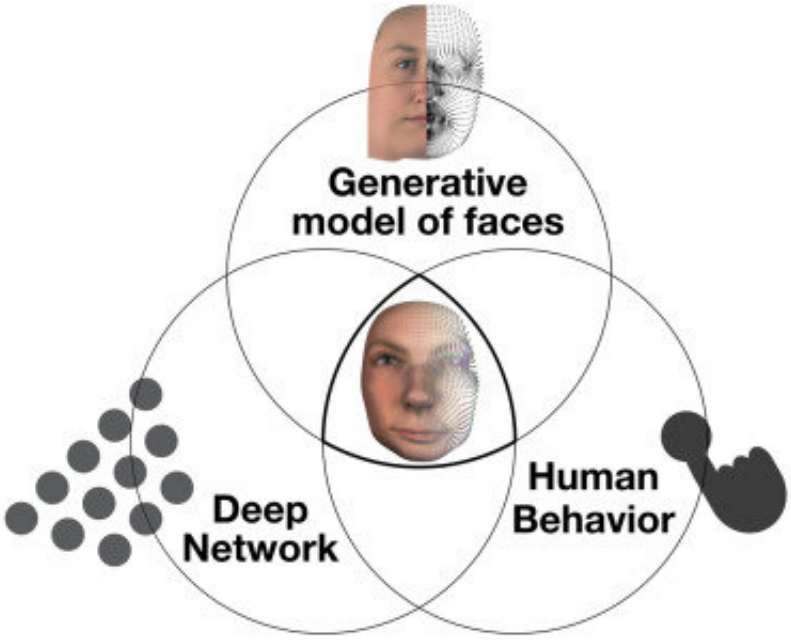
Developing an AI that 'thinks' like humans

October 11 2021



Graphical abstract. Credit: DOI: 10.1016/j.patter.2021.100348

Creating human-like AI is about more than mimicking human behavior—technology must also be able to process information, or "think" like humans if it is to be fully relied upon.

New research, published in the journal *Patterns* and led by the University of Glasgow's School of Psychology and Neuroscience, uses 3D modeling

to analyze the way Deep Neural Networks—part of the broader family of machine learning—process information, to visualize how their information processing matches that of humans.

It is hoped this new work will pave the way for the creation of more dependable AI technology that will process information like humans and make errors that we can understand and predict.

One of the challenges still facing AI development is how to better understand the process of machine thinking, and whether it matches how humans process information, in order to ensure accuracy. Deep Neural Networks are often presented as the current best model of human decision-making behavior, achieving or even exceeding human performance in some tasks. However, even deceptively simple visual discrimination tasks can reveal clear inconsistencies and errors from the AI models, when compared to humans.

Currently, Deep Neural Network technology is used in applications such a face recognition, and while it is very successful in these areas, scientists still do not fully understand how these networks process information, and therefore when errors may occur.

In this new study, the research team addressed this problem by modeling the visual stimulus that the Deep Neural Network was given, transforming it in multiple ways so they could demonstrate a similarity of recognition, via processing similar information between humans and the AI model.

Professor Philippe Schyns, senior author of the study and Head of the University of Glasgow's Institute of Neuroscience and Technology, said: "When building AI models that behave "like" humans, for instance to recognize a person's face whenever they see it as a human would do, we have to make sure that the AI model uses the same information from the

face as another human would do to recognize it. If the AI doesn't do this, we could have the illusion that the system works just like humans do, but then find it gets things wrong in some new or untested circumstances."

The researchers used a series of modifiable 3D faces, and asked humans to rate the similarity of these randomly generated faces to four familiar identities. They then used this information to test whether the Deep Neural Networks made the same ratings for the same reasons—testing not only whether humans and AI made the same decisions, but also whether it was based on the same information. Importantly, with their approach the researchers can visualize these results as the 3D faces that drive the behavior of humans and networks. For example, a network that correctly classified 2,000 identities was driven by a heavily caricaturised face, showing it identified the faces processing very different face information than humans.

Researchers hope this work will pave the way for more dependable AI technology that behaves more like humans and makes fewer unpredictable errors.

The study, "Grounding deep neural network predictions of human categorization behavior in understandable functional features: The case of face identity," is published in *Patterns*.

More information: Christoph Daube et al, Grounding deep neural network predictions of human categorization behavior in understandable functional features: The case of face identity, *Patterns* (2021). DOI: 10.1016/j.patter.2021.100348

Provided by University of Glasgow

Citation: Developing an AI that 'thinks' like humans (2021, October 11) retrieved 4 May 2024 from https://techxplore.com/news/2021-10-ai-humans.html

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.