

Merging our brains with machines won't stop the rise of the robots

27 February 2017, by Michael Milford



Merging our intelligence with that of machines has a strong appeal for some. Credit: Shutterstock

Tesla chief executive and [OpenAI](#) founder Elon Musk suggested last week that [humanity might stave off irrelevance](#) from the rise of the machines by merging with the machines and becoming cyborgs.

However, current trends in software-only artificial intelligence and [deep learning](#) technology raise serious doubts about the plausibility of this claim, especially in the long term. This doubt is not only due to hardware limitations; it is also to do with the role the [human brain](#) would play in the match-up.

Musk's thesis is straightforward: that sufficiently advanced interfaces between brain and computer will enable humans to massively augment their capabilities by being better able to leverage technologies such as machine learning and deep learning.

But the exchange goes both ways. Brain-machine interfaces may help the performance of machine learning algorithms by having humans "fill in the gaps" for tasks that the algorithms are currently bad at, like making nuanced contextual decisions.

The idea in itself is not new. [J. C. R. Licklider](#) and

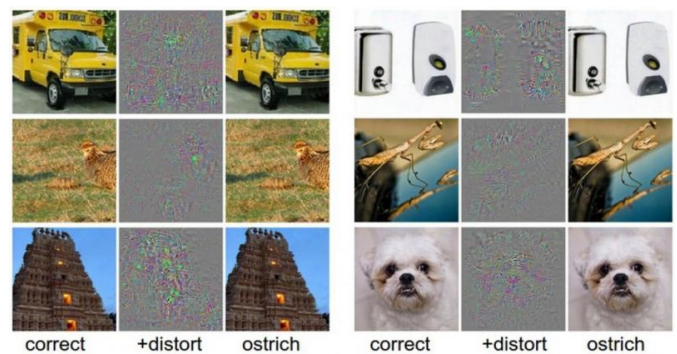
others speculated on the possibility and implications of "man-computer symbiosis" in the mid-20th century.

However, progress has been slow. One reason is development of hardware. "There is a reason they call it hardware – it is hard," [said Tony Fadell](#), creator of the iPod. And creating hardware that interfaces with organic systems is even harder.

Current technologies are primitive compared to the picture of brain-machine interfaces we're sold in science fiction movies such as [The Matrix](#).

Deep learning quirks

Assuming that the hardware challenge is eventually solved, there are bigger problems at hand. The past decade of incredible advances in deep learning research has revealed that there are some fundamental challenges to be overcome.



This image shows how you can fool AI image recognition by adding imperceptible noise to the image. Credit: Goodfellow et al, 2014

The first is simply that we still struggle to understand and characterise exactly how these complex neural network systems function.

We trust simple technology like a calculator because we know it will always do precisely what we want it to do. Errors are almost always a result of mistaken entry by the fallible human.

One vision of brain-machine augmentation would be to make us superhuman at arithmetic. So instead of pulling out a calculator or smartphone, we could think of the calculation and receive the answer instantaneously from the "assistive" machine.

Where things get tricky is if we were to try and plug into the more advanced functions offered by machine learning techniques such as deep learning.

Let's say you work in a security role at an airport and have a brain-machine augmentation that automatically scans the thousands of faces you see each day and alerts you to possible security risks.

Most machine learning systems suffer from an infamous problem whereby [a tiny change](#) in the appearance of a person or object can cause the system to catastrophically misclassify what it thinks it is looking at. Change a picture of a person by less than 1%, and the machine system might suddenly think it is looking at a bicycle.

Terrorists or criminals might exploit the different vulnerabilities of a machine to bypass security checks, [a problem that already exists](#) in online security. Humans, although limited in their own way, might not be vulnerable to such exploits.

Despite their reputation as being unemotional, [machine learning](#) technologies also suffer from bias in the same way that humans do, and can even exhibit racist behaviour if fed appropriate data. This unpredictability has major implications for how a human might plug into – and more importantly, trust – a machine.

Trust me, I'm a robot

Trust is also a two-way street. Human thought is a complex, highly dynamic activity. In this same security scenario, with a sufficiently advanced [brain-machine interface](#), how will the machine know what

human biases to ignore? After all, unconscious bias is a challenge everyone faces. What if the technology is helping you interview job candidates?

We can preview to some extent the issues of trust in a brain-machine interface by looking at how defence forces around the world are trying to address human-machine trust in an increasingly mixed human-autonomous systems battlefield.

Research into [trusted autonomous systems](#) deals with both humans trusting machines and machines trusting humans.

There is a parallel between a robot warrior making an ethical decision to ignore an unlawful order by a human and what must happen in a brain-machine interface: interpretation of the human's thoughts by the machine, while filtering fleeting thoughts and deeper unconscious biases.

In defence scenarios, the logical role for a human brain is in checking that decisions are ethical. But how will this work when the human brain is plugged into a machine that can make inferences using data at a scale that no brain can comprehend?

In the long term, the issue is whether, and how, humans will need to be involved in processes that are increasingly determined by machines. Soon [machines may make medical decisions](#) no human team can possibly fathom. What role can and should the human brain play in this process?

In some cases, the combination of automation and human workers could [increase jobs](#), but this effect is likely fleeting. Those same robots and automation systems will continue to improve, likely eventually [removing the jobs](#) they created locally.

Likewise, while humans may initially play a "useful" role in brain-machine systems, as the technology continues to improve there may be less reason to include humans in the loop at all.

The idea of maintaining humanity's relevance by integrating human brains with artificial brains is appealing. What remains to be seen is what contribution the human brain will make, especially as technology development outpaces human brain

development by a million to one.

This article was originally published on [The Conversation](#). Read the [original article](#).

Provided by The Conversation

APA citation: Merging our brains with machines won't stop the rise of the robots (2017, February 27) retrieved 20 October 2019 from <https://techxplore.com/news/2017-02-merging-brains-machines-wont-robots.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.