



AI = Artificial Imagination?

The “deepfakes” trend proves machine learning and AI have mastered the art of illusion. Are they ready to help your business innovate?

For generations, scholars and art patrons have pondered the meaning of Mona Lisa’s smile. The mystery may have recently acquired new depth, thanks to digital animations created by Samsung’s AI researchers.

Their video, viewed over one million times (and counting) on YouTube, shows the Mona Lisa **realistically moving and talking**. But this is no Hollywood CGI character created and controlled by human animators. It was achieved purely by AI – specifically, a deep neural network trained via image banks to recognise and replicate the relationships between “facial landmarks” across a range of expressions. The facial movements of the “Mona Lisa” were derived from videos of three human subjects.

Samsung’s collaboration with Leonardo da Vinci is the latest in a string of viral “deepfakes”: stunning, perhaps creepy videos in which famous faces are technologically manipulated to say and do all sorts of unlikely things. So far, the deepfake genre has given us **Barack Obama deriding Donald Trump**, and legendary surrealist artist Salvador Dali (who died in 1989) **snapping selfies** with museum-goers, among others. Of course, previous deepfake creators and their algorithms had the luxury of extensive interview footage featuring their subjects. With the Mona Lisa deepfake, Samsung has

managed to put convincing words and movements into the mouth of a single static portrait – and a painted one at that. It’s an entertaining illusion with potentially very serious implications.

How “deepfakes” work

On a basic level, the computers estimate functions of the “paths” pixels of an image should take when a human face rotates or changes expressions. These functions are estimated based on examples of movements of other human faces. Effectively the computer “learns” what a human face looks like and how it “flows” when it moves. It then “transfers” [1] this learning to any new image, like that of Mona Lisa. In doing so, it effectively generates “new numbers”, e.g. pixel values from 0 to 255, which, when rendered, result in the deepfake videos.

As surprising as it is to see a chatty Mona Lisa, the idea is, in a broad sense, the same as that of a digital calculator, introduced in the early 1960s – the mechanical precursors being centuries old. Calculators use rules provided by their developers and input data provided by the users to produce a new result (i.e. the answer to a mathematical calculation). Essentially, that is what some AI algorithms also do: They use rules that someone programmed to “generate new numbers”, in this case new pixel values. However, the major

difference is that statistical machine learning expands the field of possibilities by enabling computers to *write their own code* – which could also be their own rules – and in general “highly nonlinear functions”[2] – which auto-generates the numbers and performs tasks such as creating videos from one Mona Lisa image.

The computer is thus promoted from an assistant following the written rules of a human to an apprentice teaching itself to translate text, identify objects within images, play games... or devise realistic facial movements based on patterns gleaned from video and still images. Deepfake, then, is something of a misnomer. Strictly speaking, the visual information in these videos is no more fake than, say, **a picture that is “de-blurred”** or even a recommendation on Netflix or Amazon, or any other algorithmic data production.

Past research, in which I was lucky to participate, was in some ways a precursor of today’s Mona Lisa deepfakes. In the mid-1990s at MIT, when modern statistical machine learning was still in its infancy, my co-authors[3] and I explored using **matrix-type algebra to extrapolate images** with depth and three-dimensional structure from pairs of still photos. We were able to make our images move and talk like deepfakes, albeit not at a high enough image quality to actually fool anybody.

AI = Artificial Imagination?

At first glance, the Samsung researchers seem to have done the impossible. Technically, you would think at least two images would be needed to achieve depth and 3D structure – we need two eyes, after all. Back in the 1990s, we tried to render portraits like *Mona Lisa* as a talking head, but without a second image we were unable to do it well.[4] If machines can now create a 3D world from just one flat picture plane, can we affirm that AI possesses something akin to human ingenuity and imagination?

“Hold on”, you may be thinking. “Didn’t you just explain how machine learning operates to construct deepfakes? The ‘new, improved’ *Mona Lisa* didn’t come out of thin air or from an artist’s imagination. The computer simply learned how faces are supposed to move, transferred that knowledge to da Vinci’s masterpiece and manipulated the image accordingly.” Yet what is the big, obvious difference between a great artist’s creative skill and a bot’s preternatural dexterity with images?

What is imagination, if not the generation of new data?

So for the sake of argument, let’s grant that self-learning computers can be imaginative, according

to the above definition. It stands to reason that given enough time and technological development, they might dream up new innovations, products, etc., even without human intervention. They could learn from one domain and transfer that to other, new domains, to create innovations. It could be a major step for business, our economy and society. I would argue that this is already beginning to happen. Last year, **an AI-made artwork** sold at Christie’s auction house for nearly half a million dollars. There are countless incremental examples too – it is crucial to realise that this is *not* a black or white question. For example, when your smartphone automatically **improves upon your photography**, or autocompletes your typing with smart word suggestions, is that not a kind of creative collaboration between you and your device? Or consider that the algorithms that dispense content recommendations on YouTube and Netflix are arguably more influential than any newspaper critic. Is it crazy to speculate that these virtual curators, with their unparalleled knowledge of audience preference as well as many products, might one day not only recommend but also develop their own products and write their own screenplays?

Our work ahead

Is it at all scary that machines can now have “imagination” and be “innovative”? Yes, it can be, but this is the case with any technology. There is nothing intrinsically evil about technological development. It’s all in how you use it. As long as we maintain a safe yet flexible framework for innovation to happen, we can take advantage of amazing opportunities including in business. R&D and manufacturing, to name just two areas, could reap immense value from AI’s ability to “dream in 3D”. AI could also help researchers solve important scientific problems related to our understanding of biology, the environment, etc.

INSEAD is committed to making AI a force for good. My recent research involves deliberately diverse groups of people: not only engineers, computer scientists, business scholars and leaders, but also regulators, legal scholars, philosophers, sociologists, economists and psychologists, among others. We need a wide variety of perspectives to come to grips with the promises, as well as the threats, of new technology, particularly AI.

Happily, these days we encounter more and more senior executives who get it. They recognise that reckoning with the awesome implications of the ongoing tech revolution is a job for every level and function within the organisation. If this trend continues, features such as Facebook-style **“AI ethics boards”** and shareholder votes on controversial tech (**as recently happened at Amazon**) will become standard in every industry,

and the world will be better for it. There is simply too much at stake – and too many important and exciting things for us to do.[5]

Theodoros Evgeniou is a Professor of Decision Sciences and Technology Management at INSEAD and the Academic Director of the **INSEAD eLab**.

Found this article useful? **Subscribe** to our weekly newsletter.

Follow INSEAD Knowledge on **Twitter** and **Facebook**.

[1] There are also sub-fields in machine learning which deal with problems of so-called transfer learning and multi-task learning. These may hold the keys to major innovations in the future. We have been working on these areas at INSEAD for more than 10 years. They hold huge potential for both technology innovations and business, something we also explore.

[2] A set of rules is equivalent to a special type of functions, called “step-wise functions”. Think of an “infinite collection of rules” as becoming a general function – like a line or a parabola or any shape. See this **presentation** for more information.

[3] My co-authors for this paper were Tomaso Poggio (my PhD adviser), Amnon Shashua (co-founder of **Mobileye** and **OrCam**) and Shai Avidan.

[4] The MIT team, with my former colleague Tony Ezzat, **progressed further over the years**.

[5] The **MIT Technology Review** has just published, via the wonderful mail list “**The Algorithm**” of **Karen Hao**, a nice collection of other readings on deepfakes.

Find article at

<https://knowledge.insead.edu/blog/insead-blog/ai-artificial-imagination-11761>

Download the Knowledge app for free

