Good afternoon. Let me show you a picture of an ostrich.

This is an ostrich.

This is also an ostrich.

And this- you guessed it- is of course a picture of a frog.

Ladies and gentlemen. Welcome to the wonderful world of artificial intelligence.

First of all, I would like to make the distinction between an algorithm and an AI. Algorithms are just as they sound: lines of code, rigid and unchanging. They only do what the programmers tell them to do and nothing more. AIs, however, can actually learn from examples somewhat organically, similar to how we humans learn. And because of this, they are a lot more flexible and adaptable, often magnitudes more accurate than any handwritten algorithm. That's why they have a lot of daily applications, such as image recognition, data compression, medical diagnosis, and market prediction, just to name a few. However, unlike traditional algorithms, AIs are not ready out-of-the-box. They need to be trained, often requiring tons of training data seeing that they cannot simply draw from intuition as you and I can.

AI tasks can be generally categorised into a few problem types. The most common one being classification problems like having an AI look at emails and determine whether they are spam, or look at pictures and determine whether they are of a cat or a dog. Another common one might be generation or continuation problems where a text prompt or a half-drawn picture is given, and they are tasked with completing the rest.

Choosing the suitable AI model for a certain problem is also critical. For instance, some architectures might be better suited to natural language processing while others image synthesis. Different models are wired to give an AI the necessary tools, such as persistent memory, for problems that require it. The wrong model might not be able to tackle a certain type of problem entirely.

Now I would like to introduce you to a model called the Convolutional Neural Network, or CNN.

First question: what even is a neural network? Our brains consist of a network of neurons constantly exchanging electrical signals with each other, which is basically what "thinking" is. Neural networks for AIs are basically the same thing, but with more computer bits and less flesh and blood. It is worth noting that the majority of AIs use some form of neural network since this architecture is remarkably effective at learning and remembering.

Let’s say you’re taking a selfie with your phone. How does it distinguish your face from the random objects in the background?

That would be the task for your phone’s facial recognition CNN. CNNs are generally used for image processing tasks. They work by applying filters, or “convolutions”, to the input image over and over again until the goal is achieved.

The first filter might be concerned with finding oval shapes. This filter gets applied across the entire image, and now everything that’s oval gets highlighted. Now of course there are tons of things in the background that might trigger this first filter, and that’s why we must keep applying more filters to refine the search.

The next filter might try to find two oval shapes of roughly equal size, horizontally spaced apart from each other, and that would highlight the eyes. Now try to find another oval shape that’s below the eyes, and that would be the mouth. And so on and so forth, until it has identified all facial features and found a face with confidence.

Of course, the AI doesn’t just start with fully formed filters. In the beginning, all filters are set to random values, and it is the AI’s job to learn and tweak them during training. Each and every one of the above image filters is very much non-specific; but when combined in the right way, it can find basically any object with very high accuracy, better than humans can ever hope to.

AIs are also exceptionally great at finding patterns if there is the slightest hint of correlation. An example.

Tumours are notoriously difficult to detect in their early stages, even if doctors are given complete MRI scans of the patient at the time. It is like finding a needle in a haystack, except the needle is a tiny speck of tissue where it shouldn’t be, and the haystack is a series of greyscale blobs and shapes. But the AI is like a magnet, and it can detect such anomalies in an instant.

As a part of my group’s project, an infra-red camera is used to monitor the blood vessels in the brain. However, the captured infra-red images can be expected to be quite noisy and unclear. Fluctuations in blood pressure and blood oxygen level are undoubtedly impossible to parse conventionally. With suitable training, however, an AI can be expected to isolate the blood vessels and extract the data necessary to detect and predict strokes.

Remember when I showed you pictures of ostriches?

By adding some noise or even a single pixel onto the original image, the image recognition AI was tricked into thinking images of one thing are of another. These adversarial attacks perfectly demonstrate the exploitability and unpredictability of AI, despite its utility.

As a final word, I would like to say that we shall not forget AI is not a magical black box we can throw tasks at carelessly and expect it to solve flawlessly as we might be led to believe nowadays, or else Tesla’s self-driving cars might one day think the road is full of ostriches.

Thank you.

Generative adversarial networks (GAN)