Brain Signal AI Reads Memories and Emotions

Oskar Keding David Ohlin

What parts of your mind's contents can be read by simply looking at the electrical signals generated by the brain? Using machine learning networks trained on EEG data, we try to determine what subjects see and how it makes them feel.

Electrical signals are constantly firing across the network of neurons that make up the brain. Every conscious and unconscious process controlled by the brain gives rise to one of these signals. The *Electroencephalogram* (EEG) is recorded by attaching electrodes that measure the difference in electrical potential at points along the outside of the skull. When studying phenomena like perception and memory we try to find patterns in the EEG that correspond to relevant processes in the brain. This can be difficult due to the sheer amount of things happening in the brain at any given time, which is why we apply a number of statistical and machine learning methods to make a successful algorithm.

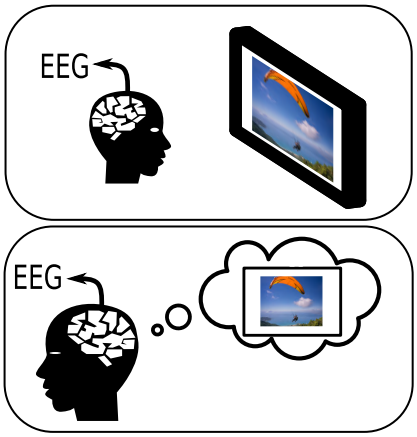

Figure 1: Outline of the experiment. First EEG is collected when the subject sees images. Then EEG is taken as the subject recollects images again.

The signals studied here originate in experiments, outlined in Figure 1, where a subject is exposed to images while sitting completely still in a controlled environment to remove as much distractions and other brain activity as possible. A second data set is generated by prompting the subject to remember the images previously shown. The images vary in both physical and emotional content and the goal of the analysis is to see which of these properties can be identified.

To make decisions about content of data one has to look at characteristics of the signals, also called *features*. An example of this is looking at frequencies, which is similar to looking at colours of a video to see what information is contained. In the study, multiple ways of extracting different features are tried. Even though we can look at these features it might be hard to find patterns that separate different categories of content from each other. We have let a computer solve this problem for us, by creating an algorithm that tries to guess the content of examples of EEG signals and compare the guess to the real content of the image. The algorithm then changes itself to make a better guess next time. This is repeated to *train* the network.

What the study found was, firstly, that it was possible for an algorithm to guess the emotional content of an image shown to a subject, by looking at the EEG. The algorithm could separate negative, neutral and positive images from each other with an accuracy significantly better than random guessing. This is intriguing, since it is then possible to teach a computer to know what makes a person feel a certain emotion.

Moreover, the study also found an interesting connection between the brain processing between the brains of different people. If one lets the algorithm *train* on a group of peoples' EEG, then the algorithm can straight away guess categories of images shown to a new person relatively accurately. Figure 2 shows this phenomenon in a plot. In the plot, the algorithm that has seen other subjects before is better with a smaller training amount. We can conclude that one person's brain, although highly personal in the way it processes information, has substantial similarities with the way other brains work.

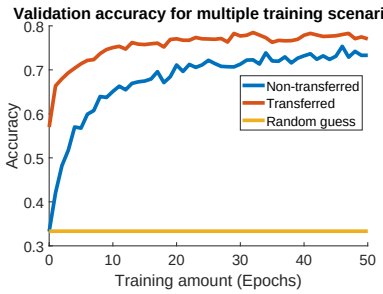

Figure 2: Performance of the model created.

In order to study the similarities between perception and memory, several tests were made with different combinations of memory and perception data. We found that algorithms trained on memories can successfully categorise data recorded during perception. This suggests that the process of remembering an image, seeing it "in your minds eye", is indeed similar to actually seeing the physical image.

As an example, aside from bringing useful insights in the ways our minds work these algorithms could, in the future, be used to enable computers to learn what you like and feel. Applications could range from less harmful things like automatic skipping of songs you don't like to more sinisterly controlling the habits of individuals. Like most areas of data science nowadays, the technologies bring lots of possible positives and negatives to the world.