

Understanding Brain Mechanisms by Integrating AI and Brain Information Analysis toward an Inclusive Society

Tomoyasu Horikawa
Distinguished Researcher, NTT
Communication Science Laboratories

Abstract

Artificial intelligence (AI) based on deep learning technologies has been appearing in many social situations in recent years, and research using these technologies has been progressing even in the research field of brain data analysis. We sat down with NTT Distinguished Researcher Tomoyasu Horikawa to talk about his research on “understanding brain mechanisms by integrating AI and brain information analysis” that aims to achieve an inclusive society by combining deep learning and brain information decoding technology.

Keywords: brain decoding, deep learning, fMRI



Achieving a deeper understanding of brain information representation by combining AI and brain information analysis

—Dr. Horikawa, please tell us about your research on “understanding brain mechanisms by integrating AI and brain information analysis.”

My aim in “understanding brain mechanisms by integrating artificial intelligence (AI) and brain information analysis” is to deepen our understanding of brain mechanisms by combining machine learning-based AI models and “brain decoding” technology. The machine learning-based AI models have come to be used in many practical situations in recent years, and deep neural network (DNN) technology, or deep learning, which has been a typical AI model, has recently come to have an even greater impact on brain

research. For a deeper understanding of the brain, I have studied means of combining such DNN technology and brain decoding, which statistically analyzes brain signals and deciphers the information represented in the brain (**Fig. 1**).

When I began my research, brain decoding technology was mainly applied to simple categorization problems with a limited set of objects (or classes) to be predicted, as in “let’s guess if the object that a person is viewing is a human face or the image of a house by analyzing brain activity measured when an image of another human face was being presented.” As technology has advanced, I have come to realize that an approach that deciphers an even greater diversity of objects could be achieved by treating the signals of an AI model (DNN), which approximates the information processing in the human brain, as the target of brain decoding.

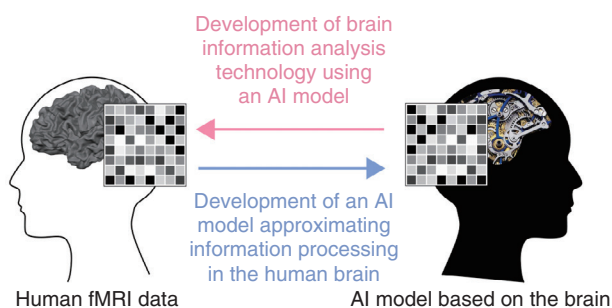


Fig. 1. Concept of understanding brain information representation using an AI model representation.

—What kind of method do you use in “understanding brain mechanisms by integrating AI and brain information analysis technology”?

The specific method that I’m using to advance my research uses data from functional magnetic resonance imaging (fMRI) that measures brain activity and a DNN model (artificial brain) that mimics the information processing in the brain. Our approach compares information representations between the human brain and the artificial brain and analyzes the human brain using the artificial brain. Based on this method, I am studying how a variety of perceptual elements are combined and represented in the brain, as in the relationship among multiple elements in a situation like “a person is riding in a car” and not simply “Is the thing that is being viewed a person or a car?” In short, in addition to individual perceptual objects in isolation, my research seeks to answer how individual elements relate to and interact with each other and how recognized information on overall meaning is represented in the brain.

In one of my previous research projects, I investigated how various types of information related to emotions are represented in the brain. A previous psychological study from another group has shown that human emotional responses could be described not only by a few categories of basic emotions such as sad or happy but also by a variety of finer categories with greater accuracy. For example, it is said that negative emotions can be classified into finer categories such as fear, horror, or disguise. Our previous study has further shown that such information related to a variety of emotions is represented in the brain based on a rating score given by a subject in response to a video. Even for this research topic, I expect that using an artificial brain, such as a DNN model, could

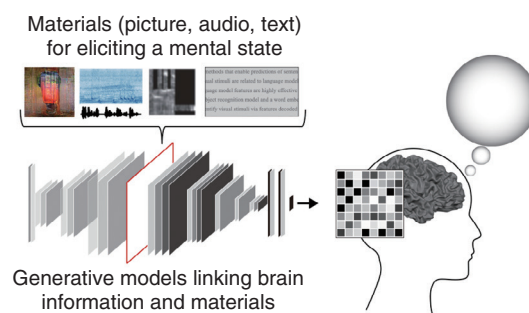


Fig. 2. Research goal: induce intended information by generating optimal stimuli based on brain information.

provide an effective means of understanding emotional representations in the brain.

The research goal shown in Fig. 2 is to generate material that induces an intended image or emotion by creating optimal pictures, audio, or text that explain specific brain states based on brain data. For example, if it can be determined how information related to emotions is represented in the brain, it may then be possible to generate images that induce targeted emotions by using brain data and DNN to generate images that intensely induce certain patterns in brain activity corresponding to those emotions.

At the laboratory that I previously belonged to, a pioneering study was being conducted on reconstructing visual images that a person was viewing based on brain data. In the early stages of that research, a visual image that was being viewed was successfully reconstructed based on brain activity data when viewing a simple pattern such as a letter or diagram as the target image. Then, in the research that I was involved in, we successfully visualized visual images for more naturalistic pictures with higher resolution using DNN technology. In past research, pictures that could be reconstructed were monochrome images 10 × 10 pixels in size, but in recent research, it has become possible to reconstruct a picture with a resolution of about 200 × 200 pixels in size reflecting detailed patterns consisting of color, object shapes, eyes or faces of animals, etc. Here, to achieve this level of accuracy, we have proved to be effective to use the knowledge that visual information is processed in a hierarchical manner in either the human brain or artificial brain (DNN). It is known that information of different levels of complexity is hierarchically represented in the brain. For example, simple visual features such as “line segments with a

certain slope” are represented in low-level regions of the brain, while features having complex shapes or patterns are represented in high-level regions. In the same way, we are coming to understand that each layer in a DNN model acquires information representations related to features that gradually increase in complexity. I believe that we have achieved such high accuracy by making use of this similarity in hierarchical representation between the human brain and artificial brain and reconstructing images using all hierarchical information from low to high levels represented in activity in the human visual cortex.

In the course of conducting such research, I am careful “not to be fooled by chance variations in the data” and to “correctly grasp what information is really represented in the brain.” In particular, as brain data are often high-dimensional and are analyzed by complex analytical procedures, there is the risk that a slight change somewhere to analysis parameters might generate what appears to be a positive result at first glance. It is certainly difficult to ascertain whether experimental data happens to show a chance positive result or reflects the true state of the brain, but we must be extremely careful not to be fooled. For this reason, I constantly ask myself, “Am I really getting correct information?” while keeping in mind the importance of correctly and faithfully carrying out the experimental plan and the procedures for data collection and statistical analysis. In addition, recent advances in AI technology enable the information decoded from the brain to be skillfully modified. For example, it is now possible to use AI to show reconstructed visual images from the brain with finer resolutions. Extracting (translating) information represented in the brain in a form that can be easily interpreted is itself important in terms of applications. However, when we consider the use of DNN with the goal of understanding the brain, it is important to cor-

rectly understand to what extent are the obtained results based on brain information and where those results have been modified by AI.

Creating an inclusive society by understanding brain mechanisms

—Please tell us your outlook for the future using “understanding brain mechanisms by integrating AI and brain information analysis technology.”

Looking to the future, I would like to expand my research in the direction of extracting information on subjective experiences represented in the human mind, developing new technologies related to my research themes such as visualization and externalization, and making practical use of technology for quantitatively evaluating information in the human brain. For example, in relatively new research conducted over these last ten years, it has been found that, when asking people to think of an apple, there are individuals who can vividly imagine a red apple with a fresh surface, whereas there are individuals that cannot imagine anything due to a condition called “aphantasia.” In my research up to now, I have been quantitatively evaluating how mental images visualized by people are represented in the brain. So using the same technique, I would like to investigate what state the brain is in when people with aphantasia attempt to generate mental imagery. Understanding the state of the brain in people without mental imagery and widely reporting our findings may help society obtain a better understanding of aphantasia and people with this condition. I hope that promoting such an understanding will lead to a more inclusive society.

—Dr. Horikawa, please leave us with a message for researchers and students.

At present, I am a member of the Sensory Representation Research Group in the Human Information Science Laboratory of NTT Communication Science Laboratories, and the full support that NTT has given me in conducting my research has been great. For example, the MRI that I use for making brain measurements is equipment that generates a powerful magnetic field, so it is necessary when presenting a subject with stimuli such as an image or audio to use a special non-magnetic device that can be used even in a high-magnetic-field environment. In addition, repeating experiments many times under a trial-and-error



process and collecting large amounts of data can incur high experimental expenses. However, over the one year since coming to NTT, the company has supported the installation of that experimental equipment and associated experimental expenses, although there had previously been no experimental environment for brain research at all in the institute. NTT has also been quite generous in procuring for me the software and hardware necessary for brain analysis. I feel that being able to pursue one's research freely and energetically in such an environment is NTT's strength. In this research environment, there are certain feelings of "joy when a paper is accepted" and "happiness when obtaining a deeper understanding of a research target." However, rather than working hard day in and day out to once again experience such joy, I feel that living a life in which I go to bed every night having thoughts like, "If I try that tomorrow, things should go well!" and wake up every morning excited about getting out to my research laboratory thinking, "OK, let's try this idea today!" is also what makes research worthwhile and even delightful. Within my long research life, I can now enjoy my research every day thanks to this NTT research environment.

On the other hand, if society is not running well, cuts may occur in the national budget for research, and it may be hard to obtain the understanding of people working in ordinary companies or of the general public with regard to research. Many other difficult situations can arise as well, such as a lack of young people who would like to become researchers or the inability of excellent researchers to continue

their research. Nevertheless, I firmly believe that research has value regardless of the time or era. Conducting research by a correct method based on a variety of ideas by many people has value, and even if there are only a few people who consider a certain research theme interesting, I think that an accumulation of research conducted by individual researchers can go on to create the world of the future. No matter how small in scale research may be, I believe that an accumulation of research projects carried out by correct methods has value. My message to everyone, and to myself as well, is "Let's do our best while believing in our own personal value!"

■ Interviewee profile

Tomoyasu Horikawa received his Ph.D. (Doctor of Science) from the Graduate School of Information Science, NARA Institute of Science and Technology in 2013. He entered NTT in 2021 and is currently a member of NTT Communication Science Laboratories. He has been a distinguished researcher since 2022. His research aims to integrate computational models inspired by the brain and brain data analysis technology. He is the recipient of the 33rd Telecommunications Advancement Foundation Award (Telecom System Technology Award) and the Best Paper Award of the Japanese Neural Network Society among other awards.