**Report on the discussion session - Session 4: AI for Neuroscience and Neuromorphic Technologies**

Yuzhe Li, OIST

Chair: Takatoshi Hikida (Osaka University)

Discussant: Kunihiko Fukushima (FLSI)

James J. DiCarlo (Massachusetts Institute of Technology)

Yukiyasu Kamitani (Kyoto University)

Rosalyn Moran (King’s College London)

Terrence Sejnowski (Salk Institute & Univesity of California San Diego)

Hidehiko Takahashi (Tokyo Medical and Dental University)

**I. Interpolating vectors**

First, Prof. Kunihiko Fukushima addressed that artificial neural network (ANN) provides design principles for brain-like AI, and is also a powerful tool for analyzing the brain. But the current AI needs big training data to learn, i.e., listen to billions, understand millions, whereas our human being learn quickly, i.e., listen to one, understand ten. Therefore, there appears to be a problem: how to let an AI learn with small data set? Prof. Fukushima proposed a method using “interpolating vectors” with a deep convolutional neural network (deep CNN). Each interpolating vector is determined on a line between two known feature vectors extracted from the training dataset, or by the nearest plane that formed by three reference vectors, or by the hyperplane that made of four feature vectors. This method achieves a data augmentation, not during the training phase, but during the prediction phase. He also proposed a computational trick to reduce the computational cost for calculating the interpolating vectors. This method achieved good learning using a small data set.

Prof. Fukushima’s talk inspired Prof. Terrence Sejnowski of the high dimensional problems that was also discussed in previous sessions. He addressed that the interpolating vector method might be an exploring tool for solving high dimensional mathematics problems.

Dr. Rosalyn Moran pointed out that using the interpolating vector as data augmentation reminds her of other data augmentation methods, such as variatiofnal auto encoder. These models are able to generate new data, so that can achieve data augmentation.

Prof. Kenji Doya asked if there is any biological interpolation of these interpolation vectors in the brain? And if there is biological evidence shows the internal data augmentation in the brain?

Prof. Fukushima answered that he believes the data augmentation should be represented somewhere in the brain, but he didn’t know where it is.

Dr. Sejnowski mentioned there are some psychological experiments showed humans can capture the center of a structure only based on the observations of some random parts, seems they are doing interpolating in the brain too. If we take fMRI while the object doing the experiment, it might give the answer where the interpolating vector is processed in the brain.

Prof. DiCarlo referred to his research in vision system. He found there are mappings between some features in the models and neurons in the brain. For example, the feature extraction is done in the high-level vision system, then can be the basis for the classifier. He thought the interpolating vectors that Prof. Fukushima proposed should be placed in the layer above the feature set, which might be in the classification layer. There are successful statistical models, but not all of them can be found in the physical groundings in the brain.

**II. Evolutionally designed V1 function**

Prof. Kenji Doya raised a question to Prof. James DiCarlo: he showed that by replacing the first layer of deep neural network by a V1-like representation, the network’s performance becomes much more robust. Then Prof. Doya’s question was: is V1 evolutionally designed, rather than trained by sensory experience, for doing the early information coding? Prof. DiCarlo commented whether the mapping comes from evolution, development, or learning is a really challenging question.

Prof. Sejnowski answered Kenji’s question: actually there exists evidence to support that V1’s function and structure are determined by evolution. There are researches that showed even before the visual system has any experience, there are neurons in V1 that become selective for orientation. Evolution is like learning, but with time step of generation. The visual cortex has more layers than in other cortex and it looks like the function and structure of V1 as the first layer seems already written in the genome. But there are more room of connection plasticity left for higher layers.

Prof. Yukiyasu Kamitani also agreed that the interpolating vector of features should be placed in a higher level. He also mentioned the price equation in population genetics. It explained how phenotype distributes across generations, which follows Bayesian formula. In this way, genome can be treated as memory, and evolution can be treated as some kind of learning. Evolution is not a random process, it has some direction.

**III. Reconstruction images from brain activity**

Prof. Fukushima asked Prof. Kamitani about his work on reconstruction of images that human subjects are looking at from their fMRI brain activities. Prof. Fukushima wondered if they can also reconstruct moving objects? Prof. Kamitani answered that they are able to decode the direction of moving dots now. And they are interested to reconstruct movies for future research. Prof. Fukushima was also interested in whether the reconstruction is possible based on the frontal cortex activities. Prof. Kamitani said the decoding from the frontal cortex activities while the subjects are sleeping also reconstructed some images, which might be the scenes the subjects were seeing during their dreams.

Prof. Moren then pointed out an interesting question: it seems that humans generate new data during dreaming, can sleeping achieve a data augmentation? Prof. Kamitani said it might be possible. Maybe there are some generative models in the brain, which might be able to achieve simulation and augmentation of new experience.

Prof. Hidehiko Takahashi added some comments to the the discussion about the high-dimensional space. From a clinical application point of view, he agreed that the clinal data are usually high-dimensional, so the methods for high-dimensional data analysis is an urgently need for their clinical use.

**IV. Universality**

Prof. Sejnowski raised a discussion on “Universality’. He reviewed some research that showed when the recurrent neural networks with different settings were trained in different tasks, even though the trained networks were overall different, the dynamics at some fixed points were similar. He said the universality is the bridge between the ANN and real neuron circuits. If the mathematical model can capture the universal concept, it would be a big breakthrough.

Kenji asked Prof. Sejnowski a question regarding converting ANNs to spiking neural networks. Even though an ANN and the real brain may use different learning algorithms, can the solution they come up be mostly the same? Prof. Sejnowski answered that is their hope. In their explorations they found the more constraints added to the deep neural network, the more architecture and the features of the brain can be captured. He also pointed out that the mathematical models can inspire people design new, more complex experiments.

Prof. DiCarlo added that this is actually what his lab is doing. They are using models to design novel experiments. Sometimes the models are shown wrong, but the experiments may be still partially correct, which leads to designing a new, more accurate model. This circle is the bridge to connect experimental direction and the theoretical direction in neuroscience.

**V. Oscillatory dynamics**

Prof. Doya asked Prof. Moran about her research using free energy framework to model the dynamics of learning and inference in the brain. Prof. Doya was curious if her model predicts any oscillatory operation in the brain.

Prof Moran said the answer is yes. In the structure of some free energy inference models, it already includes the consideration of gamma wave for forward gain control, and beta or alpha wave for backward control, which may lead to traveling waves. And there are already such applications in analysing the fMRI data and ECoG data.

Prof. Sejnowsk pointed out that there are lots of research explored representations of sensory inputs or motor outputs, but there is a lack of research on how the communications is achieved between different parts in the brain.

Maybe the oscillations is a part of the communications system, which may control the flow of information. We should be thinking about other questions.

Prof. Doya agreed with him, and said many studies assume that, when a subject performs a certain task, the required modules in the brain are activated and connected. But how such global activation and connection can be done biophysically? And what is the computational mechanism behind it? These are still big open problems.

**VI. Fixed structure but with flexibility**

Prof. Sejnowski added that the brain is not a fixed network. Each brain area is not dedicated only for one particular task. How the brain reconfigures itself for each task in a flexible way is a fascinating question.

But Prof. DiCarlo pointed out that there are a lot fixed networks in the brain. For example, the retina. Humans may overestimate the flexibility of their brains. The question should be how much flexibility we have in the brain. We can train ANN to do anything we want, even tasks humans cannot do. Most human abilities rely on the evolutionary hardware.

Prof. Sejnowski agreed with the fact that  the evolution leaves us a lot of fixed structures. But he also said there is some recent research showing that, even though the retina is well constructed, it is also adaptable, e.g., to illumination aptitudes. So there is a question raised: how could we train a fixed framework (ANN) still have the flexibility?

Prof. Churchland added that the innovations in measuring techniques and analysis changed our view of understanding the brain. Innovation in the measuring techniques revealed that neural actives reflect some changes in the environmental parameters, like Prof. Sejnowski just pointed out. On the other hand, the innovation in analysis methods for population neural actives by using cognitive models gives us a really rich picture of what the brain is doing, which is so different from what we used to focus on.