PAPER • OPEN ACCESS

A literature review of brain decoding research

View the article online for updates and enhancements.

A literature review of brain decoding research

R M Awangga*, T L R Mengko and N P Utama

School of Electrical Engineering and Informatics, Institut Teknologi Bandung, Bandung, Indonesia

*rolly@awang.ga

Abstract. Brain Decoding is a popular topic in neuroscience. The purpose is how to reconstruct an object that came from a sensory system using brain activity data. There is three brain area generally use in brain decoding research. The somatosensory area generally using mice and touch they whisker. Auditory area using different sound frequency as stimuli. The visual area using shape, random image, and video. Take one example in the visual cortex. Using the retinotopic mapping concept, the object possible to reconstruct using visual cortex activity recorded by fMRI. Retinotopic mapping focus is to relate fMRI records into visual objects seen by the subject. This brain possibilities of decoding research come to the next level when combining using deep learning. The image seen by the subject can be reconstructed by using visual cortex activity. Make reconstruction come faster and realistic to predict the stimuli. This opportunity is opening the era of the brain-computer interface. Combine a method to analyze brain functionality related to the human sensory. Bring hope and increased human quality of life. This paper reviews research in the field of brain encoding. Divide into three sections, the first section is brain decoding research in somatosensory. The second section is brain decoding in the auditory cortex. For the last section, explain visual cortex reconstruction. Every section includes equipment devices to record brain activity and the source of datasets and methods to get the brain activity data.

1. Introduction

Learn how the brain works are one of the most popular research challenges in the field of neuroscience, how the brain works in translating input from several senses. The research aims to be able to help some people who have disabilities. Some popular research conducted to study the work of the brain related to the senses, conducted on semato sensor, auditory, and visual. Focus on encoding and decoding on the human nervous system (neural) are two essential topics in the science of neuroscience cognition. Neural encoding is a representation of some stimulus features with a neuronal population as measured by neural responses. Whereas Neural decoding is a prediction of a stimulus feature with a measure of brain activity [1].

The eye is one of the five essential senses of human life. Vision is also a mandatory requirement for many productive activities and activities so that the sense of sight determines the productivity and quality of one's life where the visual function in the brain is in the occipital position called the visual cortex (Figure 1).
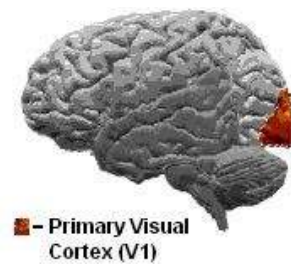
Figure 1. Primary visual cortex position (V1).

This research is a small part of the contribution to making artificial eyes. Because in general, the eye prosthesis is currently only for aesthetic purposes but does not function as a sense of sight [2]. So, from the research of image reconstruction derived from brain activity from something seen by the eye (stimulus) can contribute to the development towards the creation of a functioning artificial eye organ. To understand how the process of vision in humans, explained by the concept of the visual pathway.

1.1. Visual Pathway
The visual pathway is part of the cerebral brain that functions to process visual information. Found in occipital. Data flows from the eye through the new lateral geniculate nucleus to the visual cortex. The second figure illustrates how visual information flows.
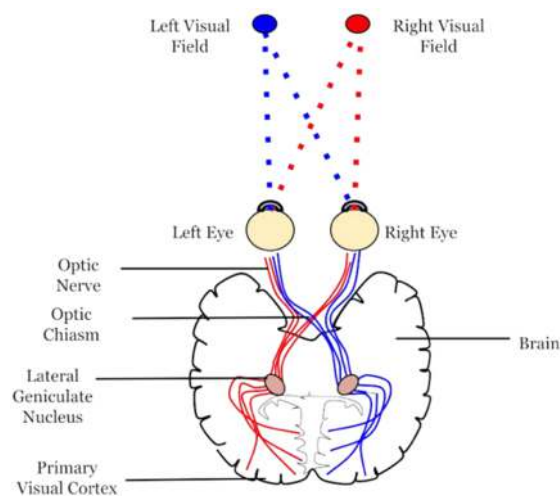

Figure 2. Visual pathway.

1.2. Types of stimuli and recording devices
In the development of pre-stimulus (*decoding*)stimulus from brain activity. In experiments that present brain activity in the visual cortex using various types of stimulation. This research uses several types of stimuli. Stimulation consists of simple shapes (geometry or numbers and letters), still and moving shading, illusion boxes (Vasarely), chess boards, colors, natural images, and moving images. In this study developed on stimulus consists of simple shapes consisting of geometry, letters, and numbers.

Several recording devices used to record brain activity in the reconstruction of stimulus image research are fMRI, MEG, EEG, and electrophysiology. The most extensive and universal recording is used for restoration and is approaching maximum results using fMRI. In the literature review section, you can see some more complete research on types of tools and examples of the stimulus.

1.3. Dataset
The study uses three datasets, namely:

- The dataset comes from the Donders Institute for Brain, Cognition, and Behavior, Radboud University Nijmegen. Get a dataset with stimulus numbers 6, and 9 [3]. The image resolution of 28 x 28. There is already a resolution between recording fMRI data from V1, V2, and V3.
- Divide data sets into two types. One type with random images, which are used to train. Another one uses geometric shapes, and the alphabet used to do the testing. The resolution used is 10x10. Uses fMRI that correlates with each image. In research conducted in the area, V1[4]. Access data at http://brainliner.jp/data/brainliner.
- The second item uses stimulus letters B, R, A, I, N, and S handwritten with the color image type greyscale. The image resolution of 56x56. In the experiment, there was a data recording on V1, and V2 [1]. The data set is at http://sciencesanne.com/research/.

1.4. Development of image reconstruction methods
Models or methods of image reconstruction that are relevant using these datasets and similar datasets can be seen in the table 1.
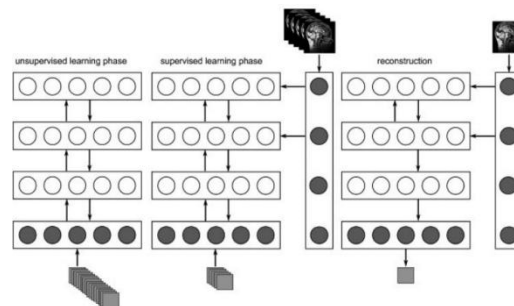

Figure 3. Illustration of the generative hierarchy model.

1.5. Research opportunities and focus
Get sure that the results of the study will be fast and precise because it uses a secondary dataset. The focus of research is the making of models with supervised learning. From several developments of image reconstruction. The development side includes:

Table 1. Model development or image reconstruction method.

Number	Model/Method	Explanation
1	Miyawaki [4]	Create a reconstruction model design using a different scale image. Use reconstruction for different stimulus scales. But one thing that is certain to be the biggest limitation is the improvement in image size, it does not apply to larger or smaller sizes. Limitations on the form of images that still produce less flexible methods. J. Mannion.
2	hierarchical generative models [3]	Use the machine concept of conditionally restricted Boltzmann by combining the methods of unsupervised learning and supervised learning. In the figure 3 contains illustrations of concepts.
3	BCCA [5]	Encoding and decoding networks are designed using a generative linear multiview model. Limitations on linear architecture and spherical covariance can hinder their performance. J. Mannion
4	Linear Recon-struction [1]	In making reconstruction patterns based on Bayesian theory using the Gaussian decoding method.

Table Cont.1.

5	Deep Co-nanically Correlated Autoencoder (DCCAE) [6]	The primary purpose of DCCAE is to do a combination of two components. These components include Canonical Correlation between two representations that are hampered by error reconstruction of each autoencoder. DCCAE can be applied using cross-view repair as neural encoding and decoding. DCCAE ignored the interview from the error reconstruction.
6	Deconvolutional Neural Network	Two layers of methods using cascade neural decoding [9]. In the first decoding, the pattern of fMRI activities brings higher feature mapping with multivariate linear regression [10]. After that, the high-level maps feature will be input for pre-trained convolutional neural networks, with outputs that can be expected to be reconstruction at each stage.
7	Deep Generative Multiview Model (DGMM) [7][8]	Combined between the Deep Generative Model to manage images received and the Speare Bayesian Linear Model to analyze brain activity. Using the principle of neural decoding, which consists of three parts, namely classification of brain activity patterns, identification of visual stimuli, and reconstruction of visual stimuli.
8	Capsule Network (CapsNet) Architecture based (CNAVR) [11]	Inspired by a mini cortical column including several million neurons in primates. The use of capsules to make groupings so that structurally organized features and representations.

(i) Comparison of current performance between algorithms.
(ii) Voxel reduction for overfitting problems.
(iii)Feature selection of activity signals in V1.
(iv)Optimization of the choice of dataset types as input from the Neural Network.

With a focus on making a neural decoding reconstruction model seen in figure 4. Building hypotheses in supervised learning, we can determine inputs and outputs. Reconstruction using data from fMRI with the target output is the stimulus used so that the goal of restoring simple stimulus can work.
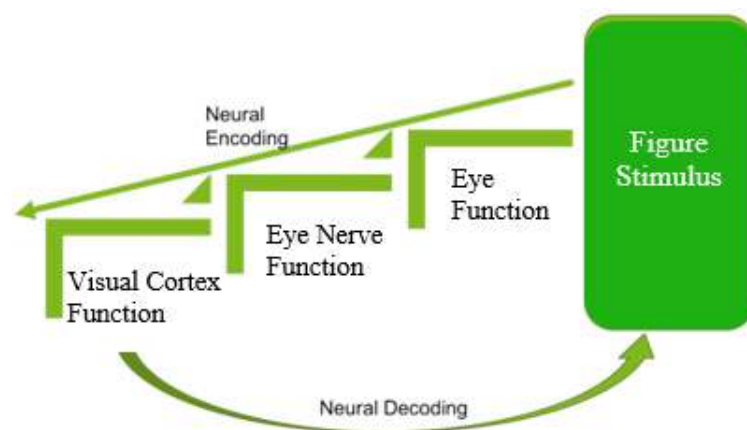

Figure 4. Block diagram of human vision function and research position.

1.6. Research stages
This research uses the methodology shown in figure 5. The first initial step is to test the dataset and compare the results between one model and another. The second stage is the trial application of RNN and the return of the model at the level of the classification algorithm. In the third stage, try to make voxel selection for efficient input determination. And in the fourth stage, a quantitative analysis of the model has been made.
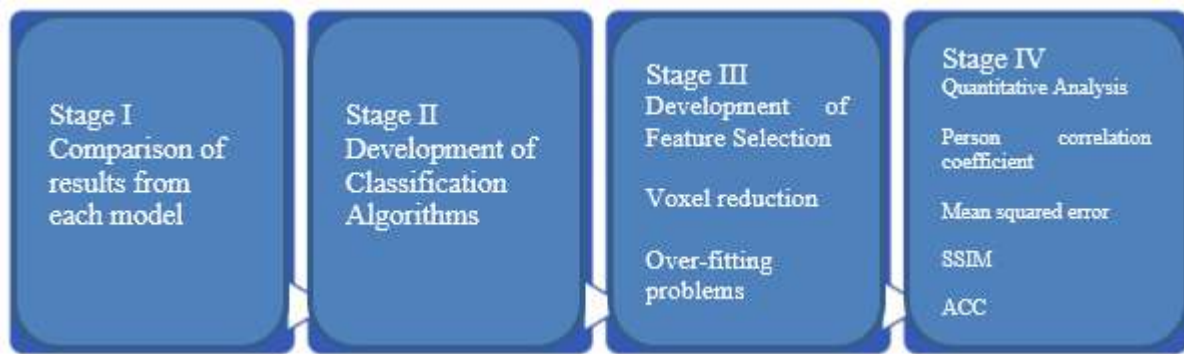
Figure 5. Research stages.

2. Related works
For data in CRCNS (Collaborative Research in Computational Neuroscience) data sharing. Among them are Somatosensory Cortex, Motor cortex, and Visual cortex. On Somatosensory Cortex There are eight datasets. On Cortex Motor There are four datasets. In the visual cortex, there are 12 PVC, one v2, three vims, 2 mt, and 1 STC. With research, subjects are mice, apes, and humans.

2.1. Visual cortex
There are 11 data available in CRCNS Primary Visual Cortex (PVC). In human research, there are fMRI data (VIM) of 3 data sets discussed in the table 2.

In the first [12] dataset, the publication by the provider of the dataset used in two publications [13,14]. In the publication [13], the problem described is on a trending topic about mind- reading, which is reading mental content from brain signals with the contribution of fMRI data set with two subject trials. Then develop a Receptive model to identify natural images from fMRI. The method used uses Pearson correlation statistics (r) and the Gabor Wavelet Pyramid. The result is a receptive model from Wavelet Gabor. It has a limitation that can only be on a pre-existing dataset for identification and is limited only to subjects that have previously tested for further testing. At the second publication in [14]. Raising the problem of mind-reading research has not yet arrived at natural images, only restricted patterns. The contribution of this paper is the use of Bayesian encoder in fMRI data to reconstruct images. By using the Bayesian Encoder and Expectation-Maximization methods. The result, a reconstruction model using a hybrid model approach consisting of structural models, semantic models, and natural images before. While the restriction of this paper is that reconstruction is only successful if the image is already in the database entered earlier, if it does not exist, then it will not work.

In the publication [15] with the [16] dataset which is the second-order dataset (vim-2). This paper presents problems with recording at Blood Oxygen Level-Dependent (BOLD) very slowly. It is not possible to reconstruct dynamic images or videos. The contribution is the reconstruction of the stimulus using dynamic images or video. By using MAP (Maximum a Posteriori) and AHP (Averaged High Reconstruction), this paper produces the Accuracy of Natural Film Reconstruction of BOLD signals by using AHP and MAP while the limitation of this paper is the absence of optimization of the reconstruction by involving energy movement models and semantic models.
In the third dataset (vim-3) paper [17,18]. Dataset describes the problem of asymmetric reconstruction between the upper and the lower limits of the stimulus shown to the subject with the contribution of the calculation of the bottom and top picture asymmetry.

Table 2. Somatosensory data.

Data Code	Data Subject	Data Explanation	Data Provider	Year	Publication	Research Opportunities
SSC-1	Mouse	Image data with fluorescent. Observe mouse that lick the left or right. Same with SSC-2, only different from its contents and completeness	Simon Peron, Jeremy Free- man, Vijay Iyer, Karel Svoboda	2014	Data explained in the paper [19], while an-other paper ex-plains the same experiment in [21]. The DOI of the dataset. Publication on paper [18]	Lick and touch mapping; lick prediction left or right
SSC-2	Mouse	Optogenetic data with a spatial resolution of 2mm and temporal resolution of 100mm. Uses a nuclear fluorescent red protein.	Simon Peron, Jeremy Free- man, Vijay Iyer, Caiying Guo, Karel Svoboda Jack L.GallantAn T.Vu, Thomas Naselaris, Yuval Benjamini, BinY Jack L.Gallant Nishimoto, An T.Vu, Thomas Naselaris, Yu- val Benjamini, BinYu, Jack L.Gallant	2015	Data explained in the paper [18], while an-other paper ex-plains the same experiment in [21]. The DOI of the dataset. Publication on paper [18]	Lick and touch mapping; lick prediction left or right, and
SSC-3	Mouse	The cortex sematori in mouse uses 512 electrodes with a distance of 60 micrometers be-tween the electrodes. 5 mi- crometer electrode diameter. Spiking activities are sorted using PCA. Data descriptions are in [23] and [24]	Shinya Ito, Fang-Chin Yeh, Nicholas M. Timme, Pawel Hot-towy, Alan M. Litke, and John M. Beggs J. Mannion	2016	Results Published on [23], [25], [22], [23],[24]	Functional neuron clustering, computational development in cortical feeds, cortical based prediction, brain connectivity mapping brain connectivity mapping prediction, brain connectivity mapping
SSC-4	Mouse	Extracellular Recording based on the movement of the mustache with specific tasks. With a touch stimulus. There is a difference in deflection speed — data in the form of brain signals.	Leah M. McGuire; Gregory Telian; Keven Laboy Juárez; Toshio Miyashita; Daniel J. Lee; Katherine A. Smith; Daniel E. Feldman	2016	Data for sup-port from the paper [20] and the dataset it-self has a DOI on [25]	Movement prediction from habits with stimulus and data clustering from movement

Table 2. Cont.

SSC-5	Mouse	Active touch to the primary somatosensory cortex (S1). Data is stored in matlab data	S. Andrew Hires, Diego Gutnisky, Jianing Yu, Karel Svoboda	2017	[26]	Movement prediction
SSC-6	Mouse	Recording 2 photon calcium in the movement of a stimulated mustache in anesthetized mouse. Ten mouse in different folders with film and stimulus documentation. Films with TIFF and stimuli are stored in the MAT file	Francisco J. Martini, Manuel Molano-Mazón and Miguel Maravall.	2017	[27]	Prediction with comparison on Visual and auditory cortex

3. Data acquisition

For data in CRCNS (Collaborative Research in Computational Neuroscience) data sharing. Among them are Somatosensory Cortex, Motor cortex, and Visual cortex. On Somatosensory Cortex There are eight datasets. On Cortex Motor There are four datasets. In the visual cortex, there are 12 PVC, one v2, three vims, 2 mt, and 1 STC. With research, subjects are mouse, apes, and humans.

3.1. Somatosensory cortex

In the Somatosensory dataset, there is a donation of data from several researchers, including the one seen in the table 3. As we know, somatosensory involves touch sensors in parts of the body, then translated into the brain.

On SSC-2 data, published on paper [18]. Data from photon imaging (figure 7) that uses red nuclear fluorescents. Experiment with observing how the mouse licked the right and left, as shown in the figure 6. Visible mapping on the results of the touch and movement of the mustache on the figure 8.

On SSC-3 data. In paper [19], wavelet transforms performed on signals at frequencies that differ from the resolution time of sub-milliseconds. Among the three brain waves, gamma and beta show the same structure between the cortex and hippocampus — a significant difference in 100-1000Hz frequency. What's interesting about the paper [19] is the map of brain connections shown in the figure 9.

SSC-4 data. The [20] paper shows how the intensity of the signal from the spiking data signal set. The figure 10 is the result of an experiment of the [20] paper with the stimulated sewer gutter condition. The impulsivity of the signals divided into three parts fast (F), medium (M), and slow (S) categories. In the part C of figure 10, the signal compressed from the right, and the left has a different pattern. The tool installed as in the figure 11.
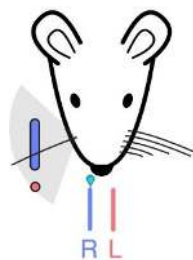

Figure 6. Experiment data recording habits of mouse lick on paper [19].

3.2. Motor cortex

The CRCNS dataset in the motor cortex consists of 4 lateral anterior motor cortex data described in the table 4.
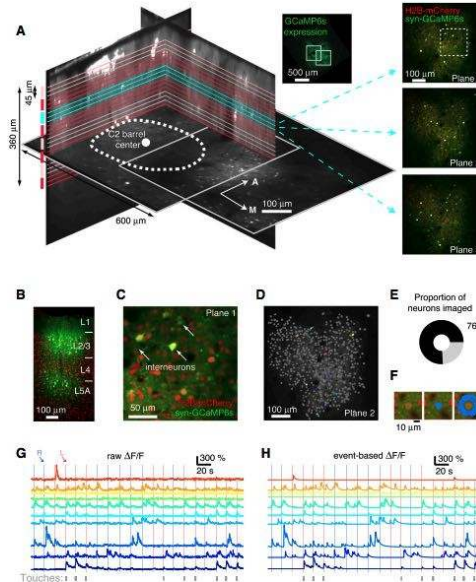

Figure 7. Photon image recording data on paper [19].

ALM-1 data uses electrophysiology to obtain electrical signals, cellular and optogenetic depictions. Description of ALM-1 data can found in paper [28]. Paper [28] shows how layer 5 of the nerve. The result is a transformation of the distribution of future activities in a movement command. This activity organized hierarchically. The figure 12 were recorded using photo inhibition. Whereas in the figure 13 recording using a silicon probe with a stimulus and using electrophysiology seen in the figure 14. Paper [28] generates a prediction of licking movements on figure 15.
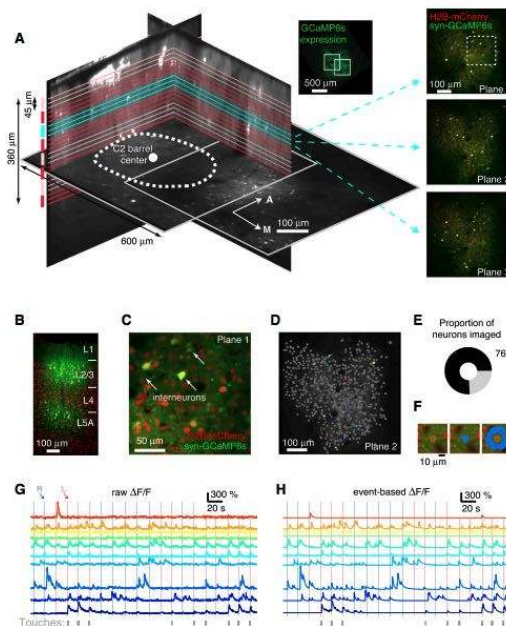
Figure 8. Map between touch and movement of a mustache in the paper [19].

From ALM-3 data with the publication [31] paper, the results explained in the form of modeling of dynamic premotor networks with ipsilateral, contralateral, and Bilateral, which can see in 16. The figure 17 is an illustrated image of how the [31] study was carried out. While the paper [32] shown the method of testing the 18 study. The [33] paper combines the Linear- Nonlinear-Poisson (NLP) model with Dynamic Time Warping (DTW) with the test picture subject in figure 19.

Table 3. Motor cortex data.

| Data Code	Data Subject	Data Explanation	Data Provider	Year	Publication	Research Opportunities
ALM-1	Adult Little Mouse	Nineteen animal objects with 99 recording is related to movement. The mouse uses its mustache to detect the two poles and then licks them (figure 12). Uses 32 channels.	Nuo Li	2014	Explanation of the data description is in the paper [28], paper that ex-plains the same experiment in the paper [21]. DOI data can access at [29]	Prediction of mouse licking movements between left and right.
ALM-2	Little Mouse	Calcium imaging of the anterior motor cortex (ALM) in the measurement of delayed touch response. The data consists of images using transient fluorescence. Data consisted of 4 subjects with 59 sessions. This data is an advanced data from ALM-1	Tsai-Wen Chen, Nuo Li, Charles R Gerfen, Zengcai V.Guo, Karel Svoboda	2016	Description of data from the results of publications at [28] with DOI dataset at [30]	Prediction of mouse movement from touch response.
ALM-3	Mouse	Nineteen subjects were mouse using electrophysiology and photo inhibition. A colostomy experiment dissected 7 of mouse	Li, Daie, Svoboda, Druckman	2016	Experiments published in paper [31] with a paper similar to the experiments on [21] which means it's still the same as the experiments in the data set ALM-1 and ALM-2	Predict the movement between lick to the left and right brain connectivity mapping prediction, brain connectivity mapping
ALM-4	Little Mouse	Thirty-four sessions on nine mouse with photo activation on fastigial and dentate. Twenty sessions on four mouse with photo inhibition.	Nuo Li	2018	Published in paper [32] DOI dataset	Searching areas in the brain for movement predictions, tuning of time, and movement parameters.
PMD-1	Monkey	Recording of cortex premotor (PMd) and cortex primary motor (M1) in sequential tasks in monkeys. The data are electrophysiological with Utah multielectrode array with two monkeys in four sessions.	Matthew G.Perich, Patrick N. Lawlor, Konrad P.Kording, Lee E.Miller	2018	Published on paper [33] with dataset [33]	Development of the use of Linear-Nonlinear-Poisson (NLP) models with Dynamic Time Warping (DTW)
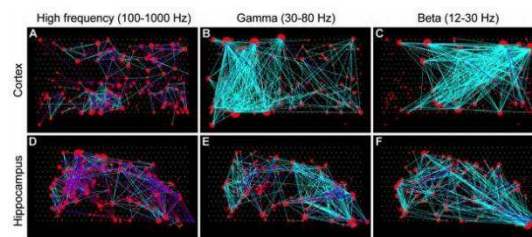
Figure 9. Map of brain connections with frequency range parameters in the paper [19].
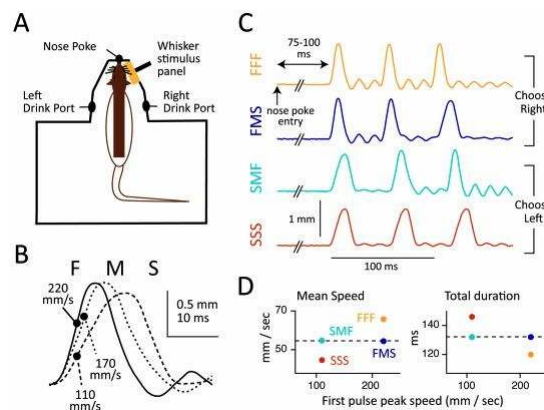

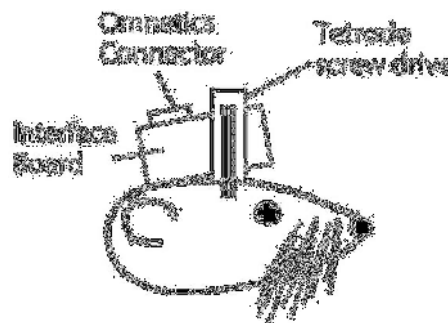**Figure 10.** Stimulus in Mouse in paper [20].

Figure 11. Installation of tools on Mouse on paper [20].

3.3. Visual cortex

There are 11 data available in CRCNS Primary Visual Cortex (PVC). In human studies, there are fMRI data (VIM) of 3 data sets described in the table 2.

In the [13] publication, the problem discussed is on a trending topic about mind-reading, which is reading mental content from brain signals with the contribution of the fMRI dataset with a trial of 2 subjects. Then develop a Receptive model to produce images of nature with fMRI. The method uses statistics Pearson (r) and the Gabor Wavelet Pyramid. Choose   a receptive model from the sports wavelet. It has a limitation that is only in the pre-existing dataset for testing, also limited to subjects that previously tested for further testing. In the publication at [14]. Lifting challenging mind-reading studies not yet on natural images is just an agreed pattern. The contribution of this paper is the use of Bayesian encoder in fMRI data to reconstruct images. By using the Bayesian Encoder and Expectation-Maximization methods. Display models using hybrid models consisting of structural models, semantic models, and previous natural images. While the limitation of this paper is that it only works if the image already exists in the previously entered database, if it's not there, then it won't work.
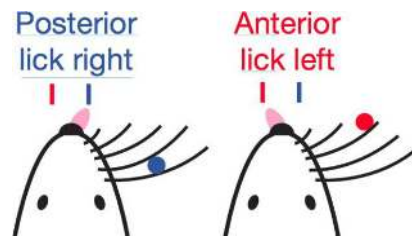
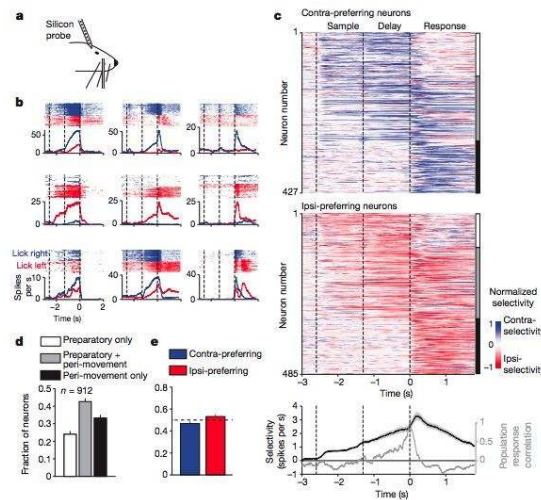**Figure 12.** Location of licking by using a mustache to detect in the paper [28].

Figure 13. ALM-1 recording with silicon probe in paper [38]

Table 4. Visual cortex data.

Data Code	Data Subject	Data Explanation	Data Provider	Year	Publication	Research Opportunities
VIM-1	Human	Human Subjects with 35 sessions, each subject separated for five days. Receptive model to identify natural images from BOLD fMRI scans	Kendrick N.Kay, Thomas Naselaris, Ryan J.	2008	On [12] and [14]	Data processing uses RCNN
VIM-2	Human	BOLD fMRI on stimulus film	Shinji Nisimoto, An T.Vu, Thomas Naselaris, Yuval Benjamini, BinYu,	2011	[15] with dataset [16]	Still using MAP and AHAP can be developed to RCNN
VIM-3	Human	fMRI in Visual cortex area v1, v2, v3	Damien J. Mannion J. Mannion	2015	[17]	Publications using GLM and spearman can trial to RCNN

In the publication [15] with the dataset which is the second-order dataset (VIM-2). This paper presents problems with recording at Blood Oxygen Level-Dependent (BOLD) very slowly. So, it is not possible to reconstruct dynamic images or videos. The contribution is the reconstruction of the stimulus using dynamic images or video. By using MAP (Maximum a Posteriori) and AHP (Averaged High Reconstruction), this paper produces the Accuracy of Natural Film Reconstruction of BOLD signals by using AHP and MAP while the limitation of this paper is the absence of optimization of the reconstruction by involving energy movement models and semantic models.

In the third dataset (VIM-3)   with paper [17]. Dataset describes the problem of asymmetric reconstruction between the upper and the lower limits of the stimulus shown to the subject with the contribution of the calculation of the bottom and top picture asymmetry. The method used is the General Linear Model (GLM) and Spearman Statistics. With the results, the effect on the bottom of the image is higher than the top of the picture. With restrictions, only discuss the asymmetry of the influence of the tip of the lower image and the upper image.
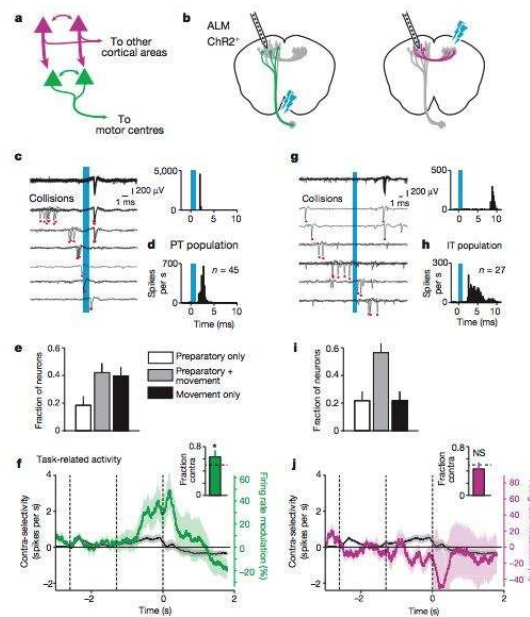
Figure 14. ALM-1 record with electrophysiology in paper [28].
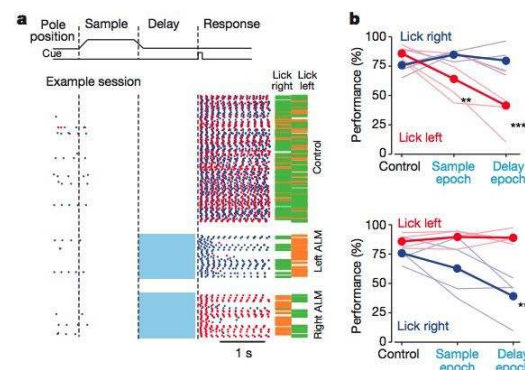
Figure 15. Calculation of motion prediction based on nerve track pyramid in paper [28].

4. Development of reconstruction research

4.1. Reconstruction using EEG

Paper [34] is a paper using an EEG tool to reconstruct the response signal. Recovery issues in this paper discuss the stimulus that drives the nerve so that it responds. Contributing contribution to this paper is to respond to the response of the EEG signal using audit and visual stimuli, including certain games and sounds with a method that combines the power of encoding and decoding. The methods and solutions used use CCA or the extension of canonical correlation analysis on the filtered stimulus called coding. Decoding itself is the response of the nerves themselves, which are related. By using EEG. As discussed in the figure 20 and the presentation of spatial data from temporal is seen in figure 21. This paper has

the main results consisting of the CCA results from the visual stimulus carried out. In the form of activity from several parts of the brain. However, this paper has the limitation of still using EEG, which is a limited tool and still has problems with the reverse problem. EEG is more temporal than spatial information. It's better to use fMRI because it's complete for personal data.
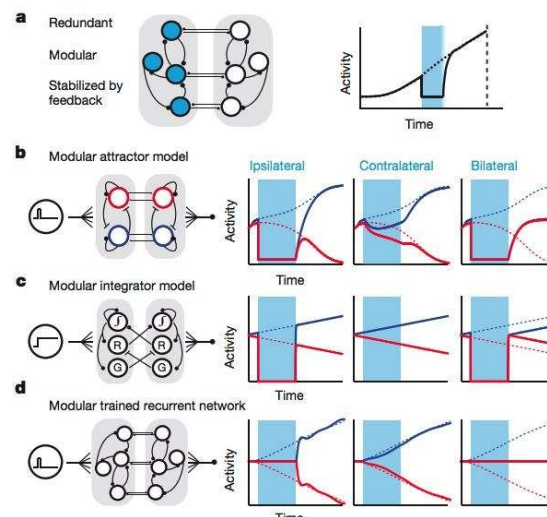

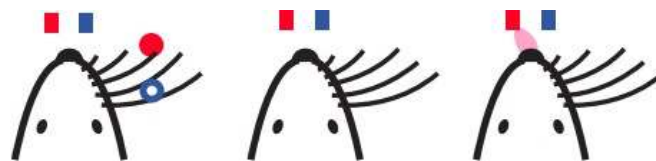**Figure 16.** Modeling motion prediction based on photo inhibition in the paper [31].

Figure 17. Illustration of how research on mouse is carried out in paper [31].
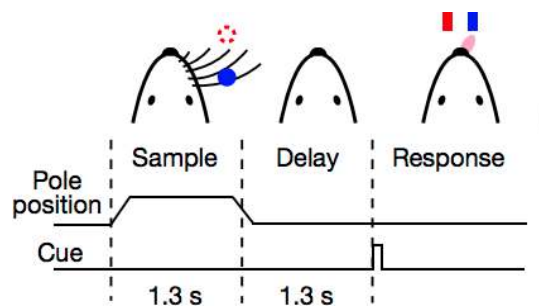
Figure 18. Illustration of how research on mouse is carried out in paper [31].
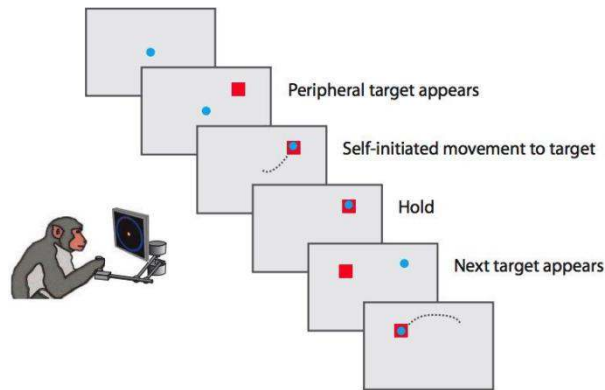
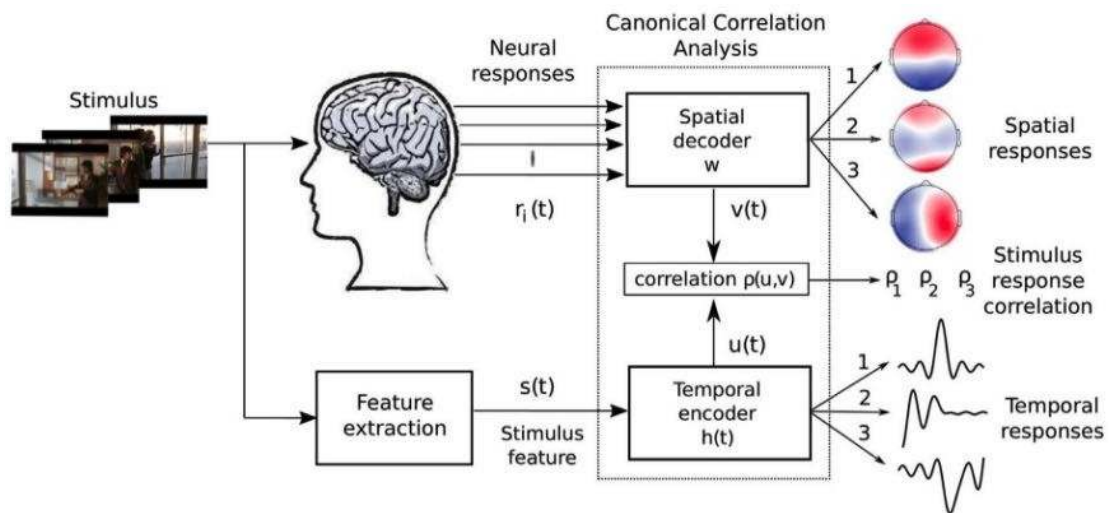**Figure 19.** Illustration of how research on monkeys is done in paper [33].

Figure 20. Spatial reconstruction using EEG.

4.2. Reconstruction using fMRI

Paper [35] regarding deconstruction of decoding of brain function. The problem in this paper is the subject of solving prediction problems with applications in the real world. Contributions in this paper explain the use of multivariate classifications before using fMRI image data processing in the brain's nerves. The methods and solutions in this paper use multivariate and univariate to predict brain function. For example, in the figure 22, multivariate is generally already used to represent objects to be represented and to be predicted. The multiplication of virgin comes from multiple voxels measured at different times, while variable predictions exist in conditions that are still experimental. Prediction or interpretation using multivariate. The difference between univariate and multivariate is seen in figure 23. In the figure 24 two experiments illustrate noise from signals using multivariate. In the figure 25 is the difference between classifications. If the different parameters differ, it will change. The main results of this paper are the differences in Univariate and multivariate decoding from the results of the analysis of the two methods before using this method for the analysis of brain function deconstruction.
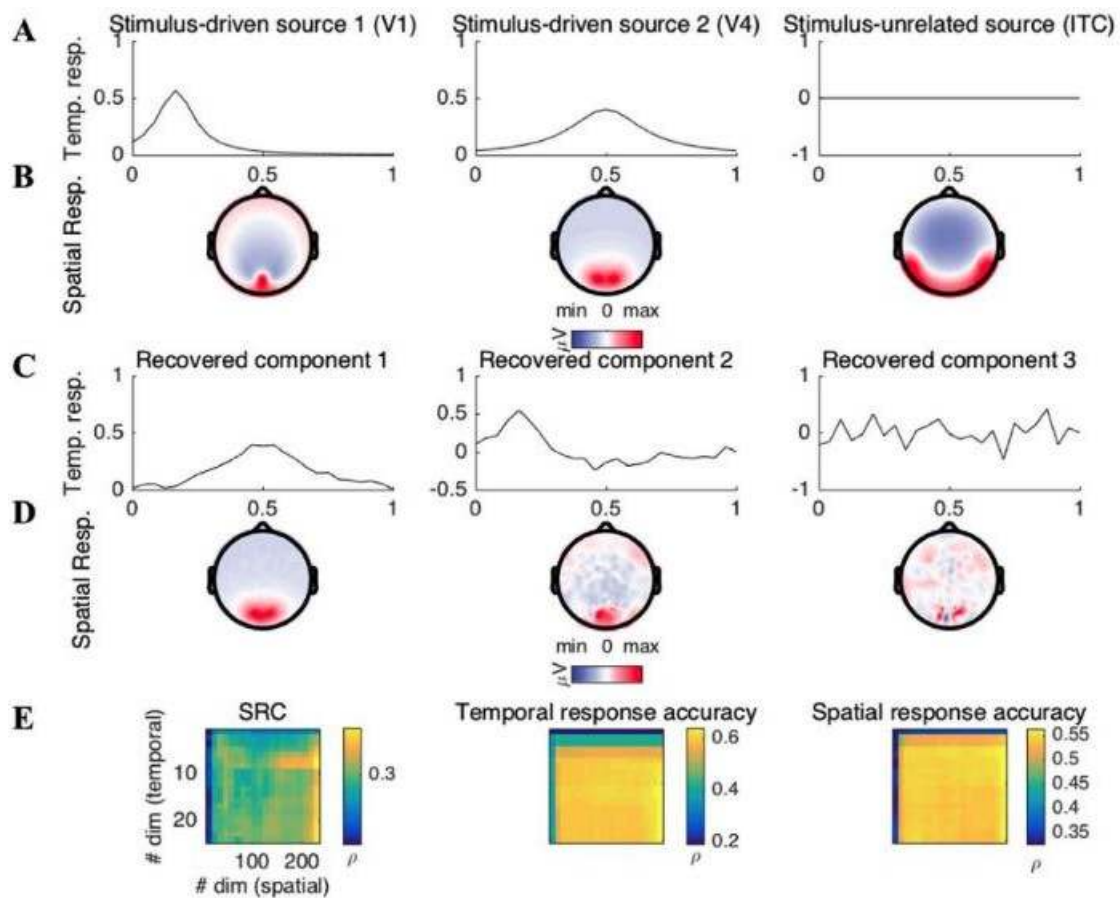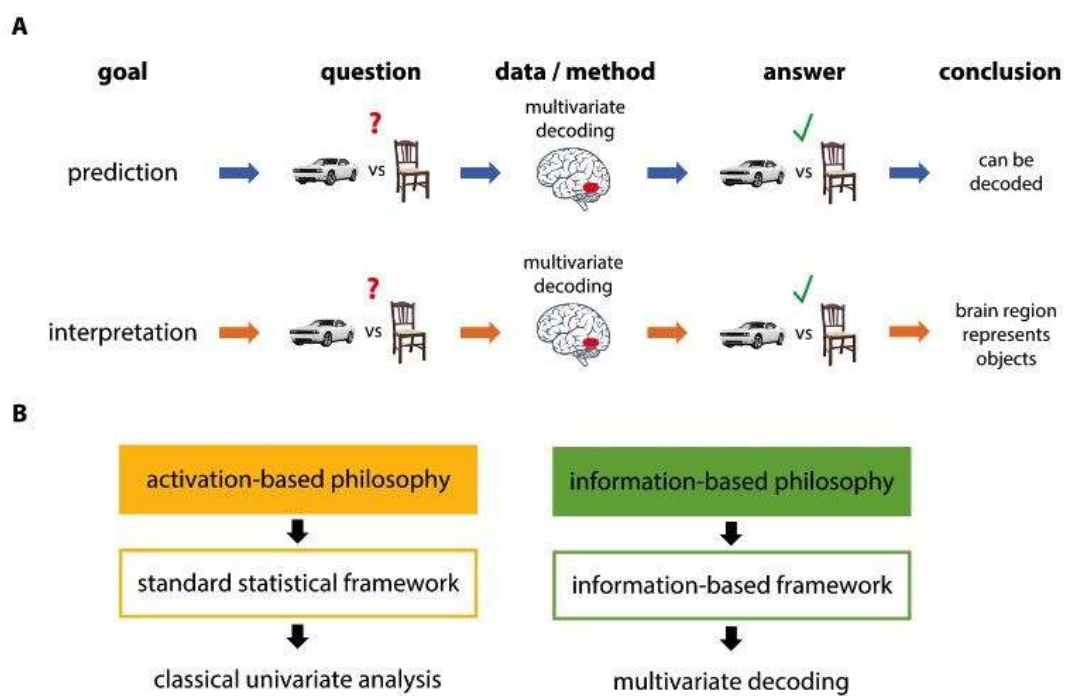
Figure 21. Spatial results from temporal EEG data.

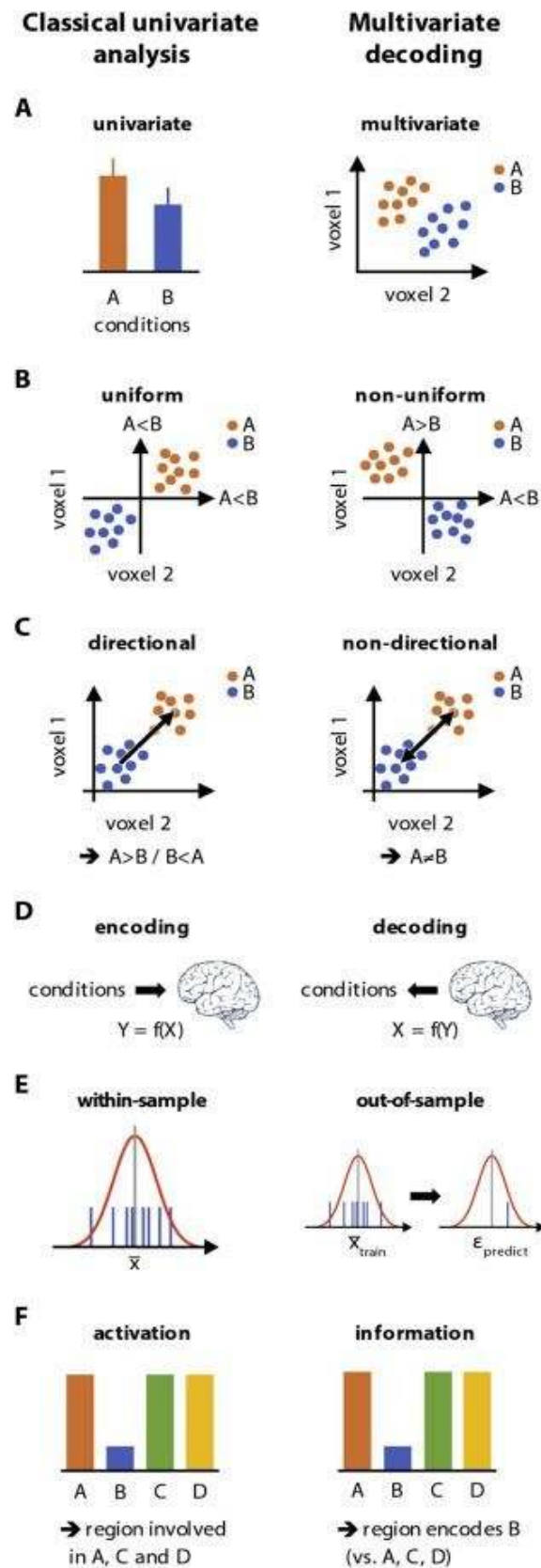**Figure 22.** The use of multivariate prediction or interpretation.

Figure 23. Difference between clusters using univariate and multivariate.
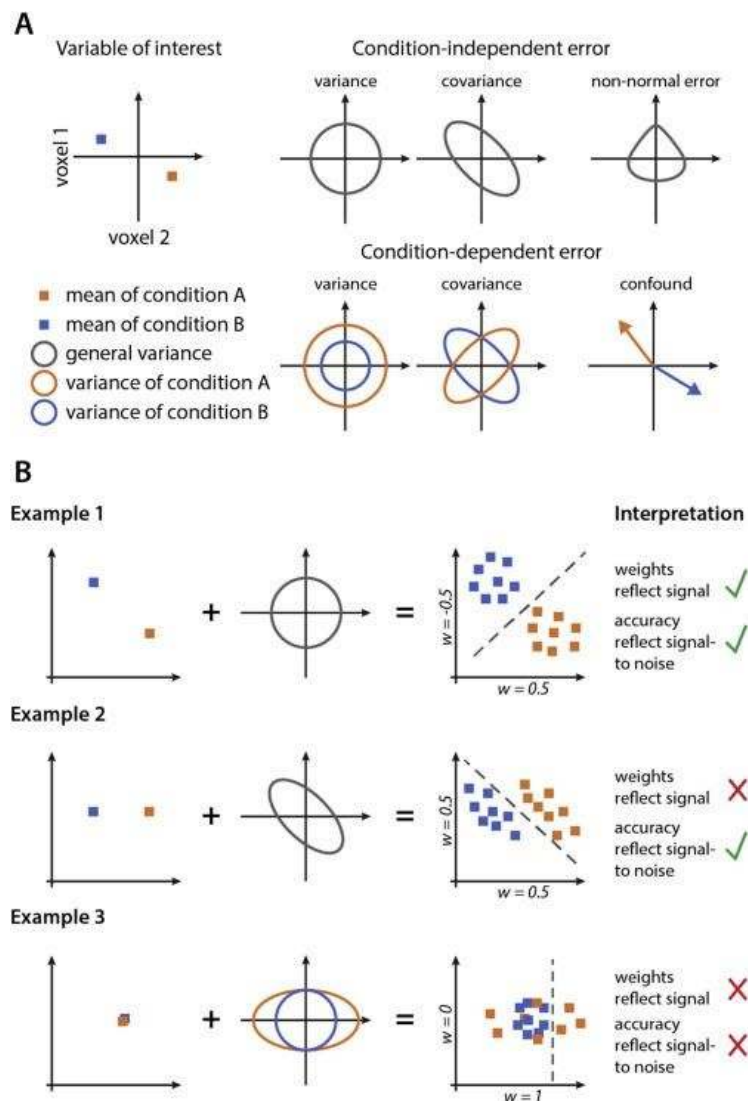
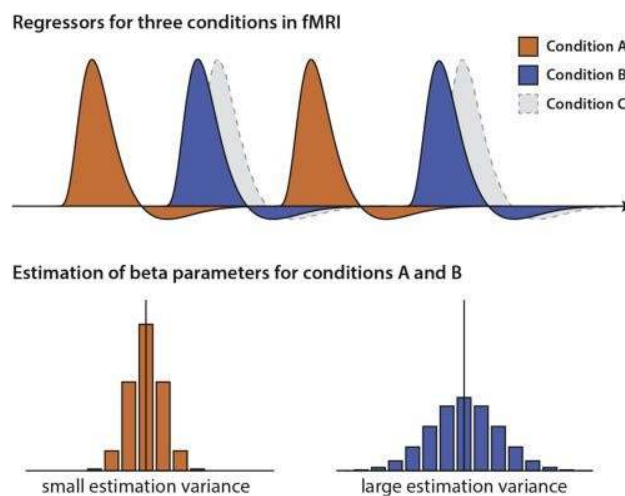**Figure 24.** Errors that occur in multivariate.

Figure 25. Classification conditions on fMRI.

Paper [36] discusses modeling patterns on voxels in fMRI. Describes the problem regarding developing Voxelwise modeling (VM), not adopting a correlation between neighbors. This paper contributes to Proposing a new framework, SPIN-VM, by adding relationships from spatial neighbors to voxels. Methods and solutions in this paper Use an experimental paradigm with a historical choreography analysis. Use five subjects for MRI in 3 sessions. The subject saw the original film then performed fMRI on BOLD. The selected voxels are estimated using VM and modeled with SPIN-VM. In a VM, each voxel is modeled independently independent of neighbors. Illustration of the picture contained in the figure 26.
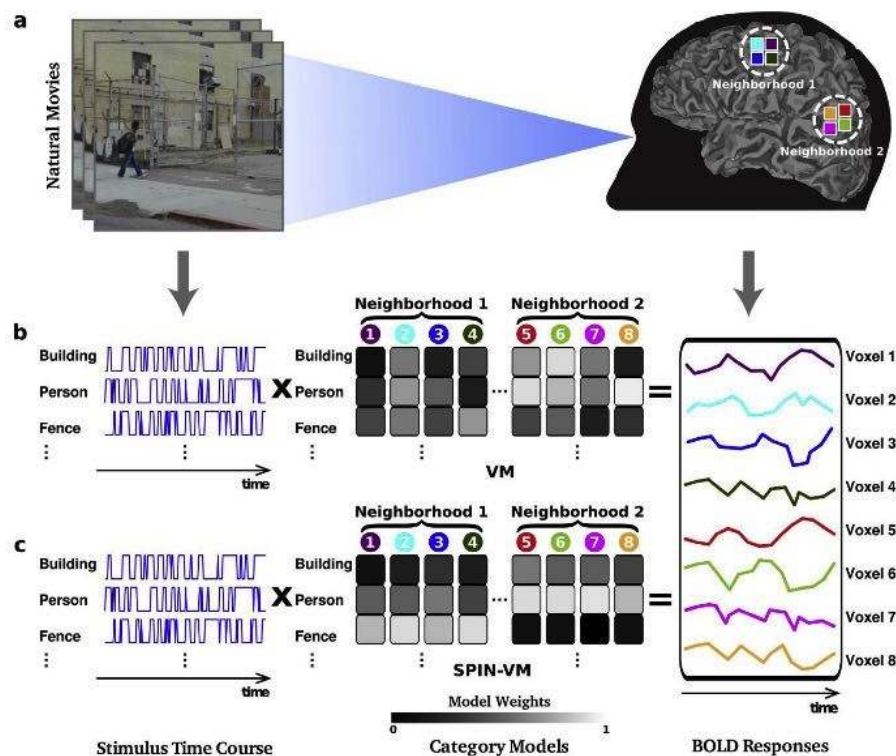

Figure 26. Illustration method.

This paper obtained the results of modeling with SPIN-VM. Selectively produces distinct object identifications. By using correlation delivers an R-value of 0.73 for SPIN-VM and 0.05 for VM. A limitation of this paper is that the SPIN-VM is compatible if the data set is limited.

4.3. Using MEG
Paper [37] uses MEG for reconstruction. This paper describes the problem of low-level dimensions of representation that do not yet have a clue in it. This paper contribution discusses developing stimulus information representation (SIR), which can eliminate information that does not come out of the cortex, which is a place to reconstruct. The method and solution use the diagnosis feature by making a stimulus into six different spatial frequencies with a circulation of 128 times. Then the perceptual stimulus and diagnosis result from the habit feature described in figure 27 is performed.
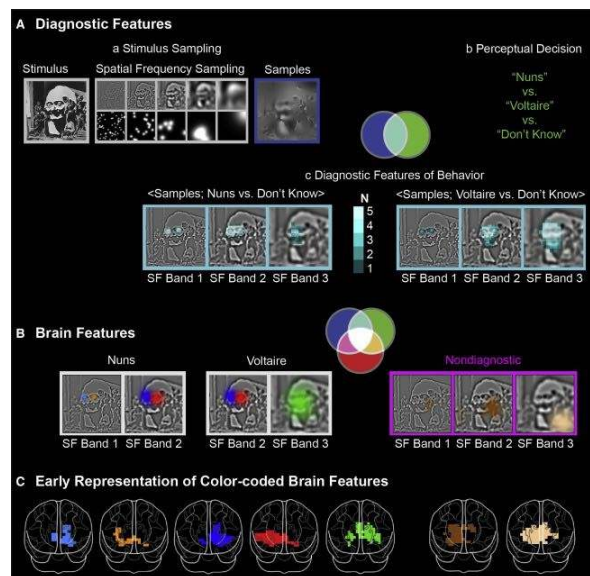
Figure 27. Illustration method.

The results of this paper are time-based and wave front based representations. Represent the location of each voxel. However, this paper has the limitation of only being able to experiment on high dimensional information.

4.4. Review paper

[38] is a review paper on computational cognition. This paper illustrates the problem of the absence of computational models that can demonstrate the habit of experimenting. Explain an overview of several scientific disciplines regarding computational cognition. The methods and solutions in this paper regarding the review paper on computational modeling of perception with the image approach from fMRI for example in the figure 28 are the calcium images in figure a, in part B the results of human fMRI using PCA (Principal Component Analysis)). Then how to know how the brain works must involve three disciplines, namely artificial intelligence, cognitive science, and computational neuroscience, as in the figure 29 to produce a computational model in the middle. In the figure 30, the position of the Deep neural network is in the middle filled with DCNN and RDNN, which begins with the V1 model wavelet Gabor. The interrelationships between the tasks, data, and models, as well as the components tested, are described in the figure 31.
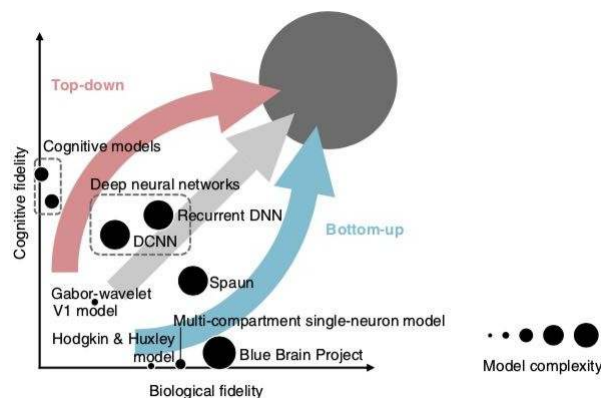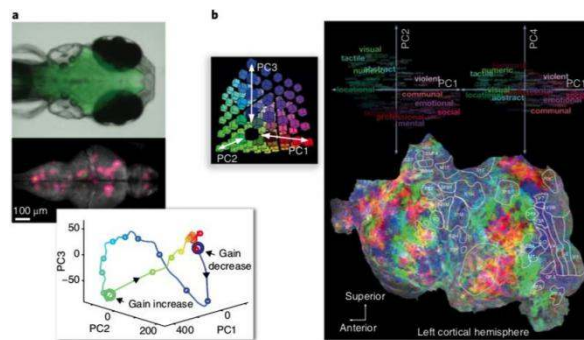

Figure 28. Neural network position.

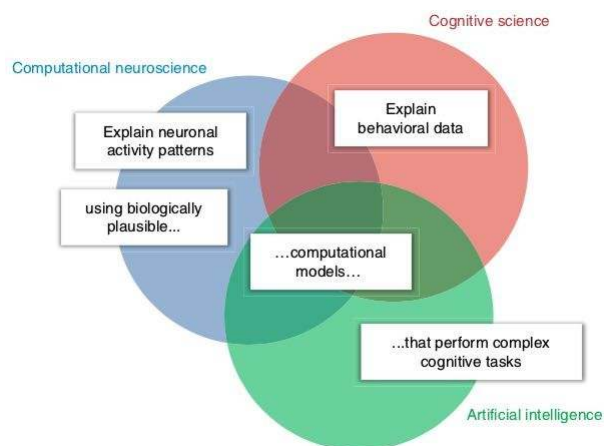**Figure 29.** Computational modelling of cognition.
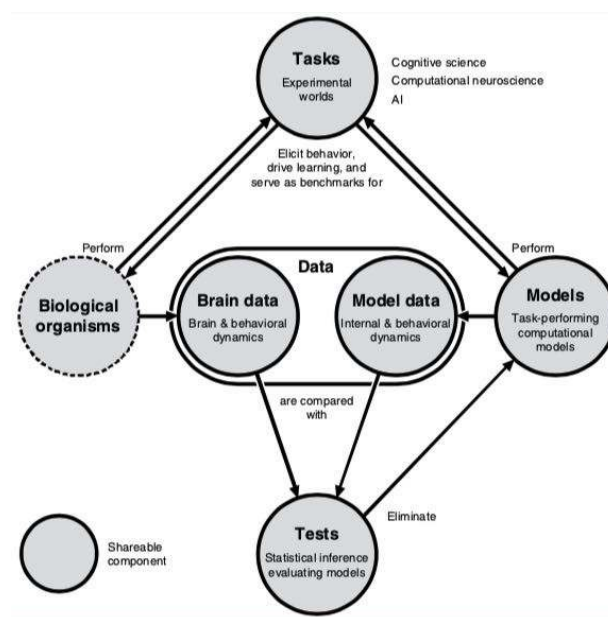

Figure 30. Three disciplines.

Figure 31. Connectivity between data.

The results of this paper have a top-down approach, and there is also a boot-up approach. This approach can have a pattern in the case of the brain, which is degenerative from the bottom up and generative at the top down. However, this paper has limitations in the review approach using the Marrs level.

References
[1]   Schoenmakers S, Barth M, Heskes T and Van Gerven M 2013 Linear reconstruction of perceived images from human brain activity *NeuroImage* **83** 951-961
[2]   Chinnery H, Thompson S B, Noroozi S and Dyer B 2017 Scoping review of the development of artificial eyes throughout the years *Edorium Journal of Disability and Rehabilitation* **3**
[3]   van Gerven M A, de Lange F P and Heskes T 2010 Neural decoding with hierarchical generative models *Neural computation* **22**(12) 3127-3142
[4]   Miyawaki Y, Uchida H, Yamashita O, Sato M A, Morito Y, Tanabe H C, and Kamitani Y 2008 Visual image reconstruction from human brain activity using a combination of multiscale local image decoders *Neuron* **60**(5) 915-929
[5]   Fujiwara Y, Miyawaki Y and Kamitani Y 2013 Modular encoding and decoding models derived from Bayesian canonical correlation analysis *Neural computation* **25**(4) 979-1005
[6]   Wang W, Arora R, Livescu K and Bilmes J 2015 On deep multi-view representation learning In *International Conference on Machine Learning* 1083-1092
[7]   Du C, Du C, Huang L and He H 2018 Reconstructing perceived images from human brain activities with Bayesian deep multiview learning *IEEE transactions on neural networks and learning systems* **30**(8) 2310-2323
[8]   Du C, Du C and He H 2017 Sharing deep generative representation for perceived image reconstruction from human brain activity In *2017 International Joint Conference on Neural Networks (IJCNN)* 1049-1056
[9]   Wen H, Shi J, Zhang Y, Lu K H, Cao J and Liu Z 2018 Neural encoding and decoding with deep learning for dynamic natural vision *Cerebral Cortex* **28**(12) 4136-4160
[10] Zeiler M D, Taylor G W and Fergus R 2011 Adaptive deconvolutional networks for mid and high level feature learning In *2011 International Conference on Computer Vision* 2018-2025
[11] Qiao K, Zhang C, Wang L, Chen J, Zeng L, Tong L and Yan B 2018 Accurate reconstruction of image stimuli from human functional magnetic resonance imaging based on the decoding model with capsule network architecture *Frontiers in neuroinformatics* **12**(62)
[12] Naselaris T, Kay K N, Nishimoto S and Gallant J L 2011 Encoding and decoding in fMRI *Neuroimage* **56**(2) 400-410
[13] Kay K N, Naselaris T, Prenger R J and Gallant J L 2008 Identifying natural images from human brain activity *Nature* **452**(7185) 352-355
[14] Naselaris T, Prenger R J, Kay K N, Oliver M and Gallant J L 2009 Bayesian reconstruction of natural images from human brain activity *Neuron* **63**(6) 902-915
[15] Nishimoto S, Vu A T, Naselaris T, Benjamini Y, Yu B and Gallant J L 2011 Reconstructing visual experiences from brain activity evoked by natural movies *Current Biology* **21**(19) 1641-1646
[16] Hanke M, Adelhöfer N, Kottke D, Iacovella V, Sengupta A, Kaule F R and Stadler J 2016 A studyforrest extension, simultaneous fMRI and eye gaze recordings during prolonged natural stimulation *Scientific data* **3**(1) 1-15
[17] Mannion D J 2015 Sensitivity to the visual field origin of natural image patches in human low-level visual cortex *PeerJ* **3** e1038
[18] Peron S P, Freeman J, Iyer V, Guo C and Svoboda K 2015 A cellular resolution map of barrel cortex activity during tactile behavior *Neuron* **86**(3) 783-799
[19] Ito S, Yeh F C, Hiolski E, Rydygier P, Gunning D E, Hottowy P and Beggs J M 2014 Large-scale, high-resolution multielectrode-array recording depicts functional network differences of cortical and hippocampal cultures *PloS one* **9**(8)
[20] McGuire L M, Telian G, Laboy-Juarez K J, Miyashita T, Lee D J, Smith K A and Feldman D E 2016 Short time-scale sensory coding in S1 during discrimination of whisker vibrotactile

      sequences *PLoS biology* **14**(8)

[21] Guo Z V, Li N, Huber D, Ophir E, Gutnisky D, Ting J T and Svoboda K 2014 Flow of cortical activity underlying a tactile decision in mice *Neuron* **81**(1) 179-194

[22] Nigam S, Shimono M, Ito S, Yeh F C, Timme N, Myroshnychenko M and Masmanidis S C 2016 Rich-club organization in effective connectivity among cortical neurons *Journal of Neuroscience* **36**(3) 670-684

[23] Shimono M and Beggs J M 2015 Functional clusters, hubs, and communities in the cortical microconnectome *Cerebral Cortex* **25**(10) 3743-3757

[24] Timme N M, Ito S, Myroshnychenko M, Nigam S, Shimono M, Yeh F C and Beggs J M 2016 High-degree neurons feed cortical computations *PLoS computational biology* **12**(5)

[25] Staff T P B 2016 Correction: Short Time-Scale Sensory Coding in S1 during Discrimination of Whisker Vibrotactile Sequences *PLoS biology* **14**(9)

[26] Hires S A, Gutnisky D A, Yu J, O'Connor D H and Svoboda K 2015 Low-noise encoding of active touch by layer 4 in the somatosensory cortex *Elife* **4** e06619

[27] Kwegyir-Afful E E, Marella S and Simons D J 2008 Response properties of mouse trigeminal ganglion neurons *Somatosensory & motor research* **25**(4) 209-221

[28] Li N, Chen T W, Guo Z V, Gerfen C R and Svoboda K 2015 A motor cortex circuit for motor planning and movement *Nature* **519**(7541) 51-56

[29] Buzsáki G and Christen Y 2016 *Micro-, meso-and macro-dynamics of the brain* (Cham: Springer)

[30] Li N, Guo Z V, Chen T W and Svoboda K 2016 *Flow of Information Underlying a Tactile Decision in Mice. In Micro-, Meso-and Macro-Dynamics of the Brain* (Springer, Cham)

[31] Li N, Daie K, Svoboda K and Druckmann S 2016 Robust neuronal dynamics in premotor cortex during motor planning *Nature* **532**(7600) 459-464

[32] Gao Z, Davis C, Thomas A M, Economo M N, Abrego A M, Svoboda K and Li N 2018 A cortico-cerebellar loop for motor planning *Nature* **563**(7729) 113-116

[33] Lawlor P N, Perich M G, Miller L E and Kording K P 2018 Linear-nonlinear-time-warp-poisson models of neural activity *Journal of computational neuroscience* **45**(3) 173-191

[34] Dmochowski J P, Ki J J, DeGuzman P, Sajda P and Parra L C 2018 Extracting multidimensional stimulus-response correlations using hybrid encoding-decoding of neural activity *NeuroImage* **180** 134-146

[35] Hebart M N and Baker C I 2018 Deconstructing multivariate decoding for the study of brain function *Neuroimage* **180** 4-18

[36] Çelik E, Dar S U H, Yılmaz Ö, Keleş Ü and Çukur T 2019 Spatially informed voxelwise modeling for naturalistic fMRI experiments *NeuroImage* **186** 741-757

[37] Ince R A, Van Rijsbergen N J, Thut G, Rousselet G A, Gross J, Panzeri S and Schyns P G 2015 Tracing the flow of perceptual features in an algorithmic brain network *Scientific reports* **5** 17681

[38] Kriegeskorte N and Douglas P K 2018 Cognitive computational neuroscience *Nature neuroscience* **21**(9) 1148-1160