



HAL
open science

Exploring the Relationship Between Visual Information and Language Semantic Concept in the Human Brain

Haodong Jing, Ming Du, Yongqiang Ma, Nanning Zheng

► **To cite this version:**

Haodong Jing, Ming Du, Yongqiang Ma, Nanning Zheng. Exploring the Relationship Between Visual Information and Language Semantic Concept in the Human Brain. 18th IFIP International Conference on Artificial Intelligence Applications and Innovations (AIAI), Jun 2022, Hersonissos, Greece. pp.394-406, 10.1007/978-3-031-08333-4_32 . hal-04317198

HAL Id: hal-04317198

<https://inria.hal.science/hal-04317198v1>

Submitted on 1 Dec 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



This document is the original author manuscript of a paper submitted to an IFIP conference proceedings or other IFIP publication by Springer Nature. As such, there may be some differences in the official published version of the paper. Such differences, if any, are usually due to reformatting during preparation for publication or minor corrections made by the author(s) during final proofreading of the publication manuscript.

Exploring the Relationship Between Visual Information and Language Semantic Concept in the Human Brain^{*}

Haodong Jing¹², Ming Du¹², Yongqiang Ma¹² and Nanning Zheng¹²³

¹ Institute of Artificial Intelligence and Robotics,
Xi'an Jiaotong University, Xi'an, Shaanxi, P.R.China

² National Engineering Laboratory for Visual Information Processing and Applications,
Xi'an Jiaotong University, Xi'an, Shaanxi, P.R.China

{ jinghd , x1978626209 }@stu.xjtu.edu.cn;
musayq@xjtu.edu.cn

³ Correspondence: nnzheng@mail.xjtu.edu.cn

Abstract. Functional magnetic resonance imaging (fMRI) can be used to map patterns of brain activity and understand how information is expressed in the human brain. Using fMRI data to analyse the relationship between visual cortex and language semantic representation is of significance for building a new deep learning model. The cognition of visual semantic concept refers to the behavior that people can distinguish and classify the semantic concept of visual information they see. Many previous research literatures have revealed semantically active brain regions, but lack of modeling the relationship between visual information and language semantic concept in human brain, which is very important to understand the brain mechanism of concept learning. In this paper, we propose a Semantic Concept Cognitive Network (S-ConceptNet) model of brain cortical signals based on fMRI. The model organizes visual and linguistic semantic information into a unified representation framework, which can effectively analyse the generation process of semantic concept, and realize the function of semantic concept cognition. Based on S-ConceptNet, we also use the Dual-learning model to reconstruct the brain signal, judge the corresponding concept category through the S-ConceptNet, compare the reconstructed image with the image of this category. And finally output semantic information corresponding to the brain signal through similarity. We verify the effect of the model and conduct comparative experiments, and the experimental results are better than previous work, and prove the effectiveness of the model proposed in this paper.

Keywords: Brain-inspired computing, Semantic perception, fMRI, Semantic Concept Cognitive Network, Dual-learning

1 Introduction

In recent years, the research of artificial intelligence has made great progress, but the current artificial intelligence system is still far from the function of the human brain [1]. More and more researchers realize that only by fully understanding the brain can

^{*}This work was supported by the National Natural Science Foundation of China (NO. 62088102) and China National Postdoctoral Program for Innovative Talents from China Postdoctoral Science Foundation (NO. BX2021239)

machines be able to think, reason and learn flexibly and quickly [2]. In order to achieve this goal, in recent years, more and more scientists have invested in artificial intelligence research based on neuroscience (especially the research of brain cognitive mechanism and the modeling of brain function), and constructed new artificial intelligence from cognitive processes [3]. AI algorithm framework. The analysis of fMRI brain signals under visual stimuli is a research hotspot in this field [4].

In past, a series of studies have been carried out on fMRI visual brain signals under semantic stimuli using various methods. Initially, people used statistical methods and simple linear models to model fMRI brain signals, and used Bayesian inference [5, 6] to analyze the data. In recent years, deep neural networks (DNNs)[7], VAEs [8], DCCA [9], semantic attractor networks [10] and other methods [11, 12] have been used to study this issue. In the research on how the brain comprehends semantic textual information, in the past ten years, most researches have focused on the study of "brain signals under single semantic stimulation", or the cognitive process of semantic brain signals in English. However, there is a lack of relevant research and analysis on the cognitive neural basis of Chinese semantic processing and the understanding of semantic concepts, especially the understanding of semantic concepts under visual conditions.

The current research on visual cognition and semantic understanding of the brain includes primary visual cortex and higher visual cortex (HVC), left inferior frontal gyrus (IFG), medial prefrontal cortex (mPFC) [13, 14, 15], The information of these brain regions can reveal the internal information flow of the brain in the process of semantic cognition, which can help interpret the operation principle of the human brain [16, 17].

Our team has been committed to exploring brain signal related research. The motivation of our research is that the understanding and classification of semantics is very important in the research of artificial intelligence, which can help machines to reason and understand the world more intuitively, rather than the mechanized and poorly generalized classification generated after a lot of training. model (like some visual recognition tasks), such human brain research and corresponding model building is meaningful for the further development of artificial intelligence.

Previously, we proposed related models such as the brain signal analysis model based on voxel correlation. On the basis of previous work, we consider how to introduce the understanding of semantic concept information in the visual cognitive process of the brain based on the correlation analysis model, so as to improve the understanding of the brain and information reconstruction.

The Semantic Concept Cognitive Network model (S-ConceptNet) proposed in this paper uses the powerful function of recurrent neural network to fit the cognitive function of the brain to classify concepts and semantic feature values in our cognitive experiments. The cognitive function of semantic concepts is then based on the S-ConceptNet model, this paper establishes the image reconstruction and concept classification model of brain signal, so as to verify the validity of the model.

2 An End-to-End Dual-Learning Method

Brain decoding based on fMRI brain signals usually requires the acquisition of a large number of pairs of (Image, fMRI) data. However, in many studies, it is difficult

to obtain a large number of pairs of data, and limited data has many limitations in the training process of the model [18].

In recent related research, some scholars have proposed a new idea of Dual-Learning to solve related problems. It can overcome the inherent lack of training data by introducing self-supervision using unlabeled data. The Dual-learning trains two types of networks: Encoder, which encodes the Image to the corresponding fMRI data; Decoder, which decodes the fMRI to the corresponding Image data. Cascading the two networks forward, The En-De network produces a model whose input and output are the same Image, which enables unsupervised training on a large number of images recorded with brainless signals. Then, concatenating the two networks in reverse, De-En, can produce a model whose input and output are fMRI signals and have the same weights as the previously trained network. The encoding and decoding models represent the original bidirectional dual pair and form a closed loop, allowing the application of bidirectional learning and unsupervised training on fMRI data [19, 20, 21].

But with further research, we believe that the previous work could be improved:

First, the bidirectional learning model does not have sufficient application of labeled data pairs, and only uses a large number of unpaired fMRI and image data to train the encoding and decoding process. There is room for improvement in data application and loss function definition.

Second, the image restoration model constructed by Dual-Learning only stays on the statistical features of voxel data without further analysis of the internal relationship of voxel information. It is essentially a pixel-based image processing method, which cannot be effectively used. To understand semantic information and classify concepts, The fit to the brain's processing of semantic conceptual information can be improved.

3 Semantic Concept Cognitive Network

One of the most important challenges in semantic concept cognition is how to describe the correspondence between semantic concepts and acquired fMRI data. To characterize hidden associations between two different forms of data, we propose a cross-modal ranking loss for semantic stimulation and reconstructed brain signals. The corresponding schematic ordering in semantic concept stimulus classification is shown in Fig.1.

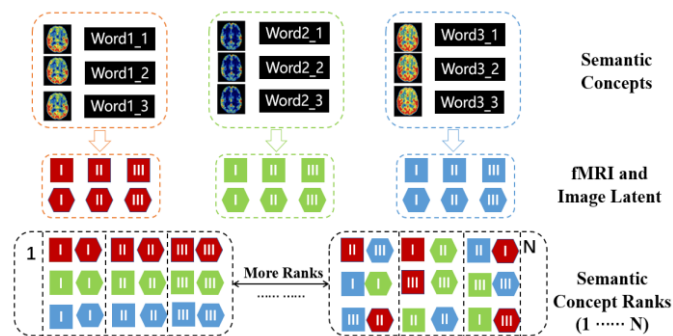


Fig. 1. Ranking Loss in Semantic Concept Understanding

In accomplishing such a ranking loss, we consider an LSTM network for fMRI feature ranking: a common LSTM unit is composed of a cell, an input gate, an output gate and a forget gate. The cell remembers values over arbitrary time intervals and the three gates regulate the flow of information into and out of the cell.

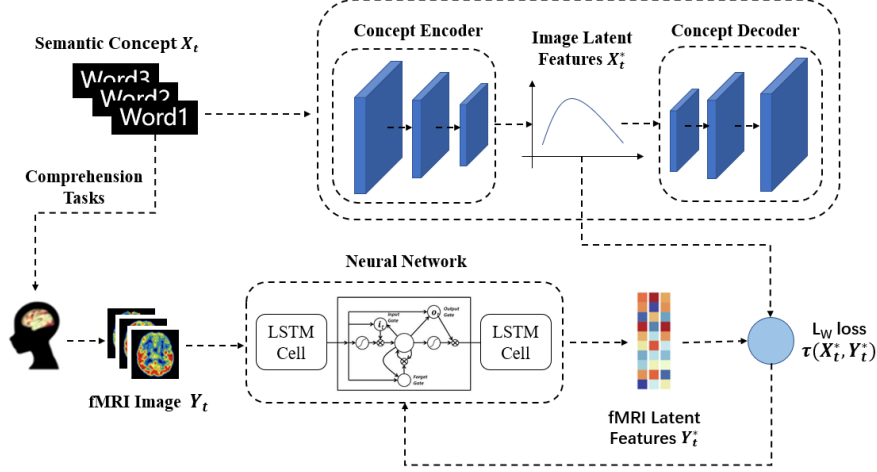


Fig. 2. The framework of Semantic Concept Cognitive Network.

The main components of the semantic concept cognitive model constructed in this paper are shown in the Fig.2. We respectively construct a network that extracts eigenvalues from the stimulus image and from the fMRI brain signal for the construction of the semantic concept cognitive model.

Starting from the stimulus image, we characterize the feature values of the semantic concept image through the corresponding encoder deep network, and map it into the image latent space z^i . The LSTM network is used to extract the latent eigenvalues of brain signals and map them into the fMRI latent space z^f .

Here, we denote (X_t^*, Y_t^*) as the semantic image and fMRI brain signal at time point t generated by the stimulus image encoder. We further use $(X_t^{*'}, Y_t^{*'})$ to denote non-corresponding samples, where $X_t^{*'}$ denotes stimulus images independent of $Y_t^{*'}$, and $Y_t^{*'}$ indicates brain activity not induced by $X_t^{*'}$. Our model ensures that fMRI combinations with high correlations are in the front and those with low correlations are in the tail. Specifically, for the semantic information of the same concept, it will be given the highest ranking, for a pair of the same concept but not the correct category, it will be set as medium, and for the wrong concept identification, it will be classified as an irrelevant sample, and given The lowest weight for this sample. We give the ranking loss as follows:

$$L_{Loss} = \lambda_w L_W = \frac{1}{T} \sum_{t=1}^T L_{Weight}(X_t^*, Y_t^*) \quad (1)$$

Among them, the ranking loss $L_{Weight}(X_t^*, Y_t^*)$ of a single paired sample is defined as:

$$L_{Wight} = \sum_{X_t^*} [\gamma - \tau(X_t^*, Y_t^*) + \tau(X_t^*, Y_t^*)]_+ + \sum_{Y_t^{*'}} [\gamma - \tau(X_t^*, Y_t^{*'}) + \tau(X_t^*, Y_t^{*'})]_+ \quad (2)$$

In the formula, γ is the margin, $\tau(X_t^*, Y_t^*) = -\|(\max(0, X_t^* - Y_t^*))\|^2$ is used to measure the stimulus image and evoked brain Order violation penalties for similarity between activities. Also, $[x]_+$ means $\max(x, 0)$

λ_W is a hyperparameter that balances the influence of the loss function on the overall model.

In the network eigenvalue matching training, using the eigenvalue correlation ranking mentioned above, the latent space of the two types of data is mapped into a common space through the sorting loss L_W , and the corresponding eigenvalues of different concepts are carried out. The weights are sorted so that the final trained network can form a corresponding pairing relationship between the semantic images of the same conceptual category and the fMRI brain signals.

In the process of model application, in the testing phase, we input the fMRI brain signals collected by the experiment, the model will label different concept categories, and identify the concept category to which the input brain signal belongs, and our semantic concept category represented by the brain signal, which helps us further understand semantic information.

The algorithm of Semantic Concept Cognitive Network is shown in Algorithm.1.

Algorithm 1 Semantic Concept Model

Require:

Semantic images and fMRI brain signals at n time points t generated by the stimulus image encoder $(X_t^*, Y_t^*) = \{(X_1^*, Y_1^*), (X_2^*, Y_2^*), \dots, (X_n^*, Y_n^*)\}$, m non-corresponding samples $(X_t^{*'}, Y_t^{*'}) = \{(X_1^{*'}, Y_1^{*'}), (X_2^{*'}, Y_2^{*'}), \dots, (X_m^{*'}, Y_m^{*'})\}$, where $X_t^{*'}$ represents stimulus images independent of $Y_t^{*'}$, while $Y_t^{*'}$ represents brain activity not evoked by $X_t^{*'}$

Ensure:

min L_{loss} ; **The best sorting result**

- 1: **First, The governing equations of LSTM are defined as follows:**
- 2: $i_t = \sigma(y_t + U_i \hat{h}_{t-1} + b_i)$
- 3: $f_t = \sigma(y_t + U_f \hat{h}_{t-1} + b_f)$
- 4: $\tilde{c}_t = \tanh(W_c y_t + U_c \hat{h}_{t-1} + b_c)$
- 5: $c_t = i_t * \tilde{c}_t + f_t * c_{t-1}$
- 6: $o_t = \sigma(W_o y_t + U_o \hat{h}_{t-1} + b_o)$
- 7: $\hat{h}_t = o_t * \tanh(c_t)$
- 8: **for** $x = 1$ to n **do**
- 9: **for** $y = 1$ to m **do**
- 10: $L_{Wight} = \sum_{X_t^*} [\gamma - \tau(X_t^*, Y_t^*) + \tau(X_t^*, Y_t^*)]_+ +$
- 11: $\sum_{Y_t^{*'}} [\gamma - \tau(X_t^*, Y_t^{*'}) + \tau(X_t^*, Y_t^{*'})]_+$
- 12: $\tau(X_t^*, Y_t^*) = -\|(\max(0, X_t^* - Y_t^*))\|^2$
- 13: $L_{loss} = \lambda_W L_W$
- 14: **update** L_{Wight}
- 15: **end for**

- 16: $L_W = \frac{1}{T} \sum_{t=1}^T L_{Weight}(X_t^*, Y_t^*)$
 17: **update** c_t, \hat{k}_t
 18: **end for**

4 Brain Signal Semantic Understanding Model

The model constructed in this paper is shown in the following figure.3.

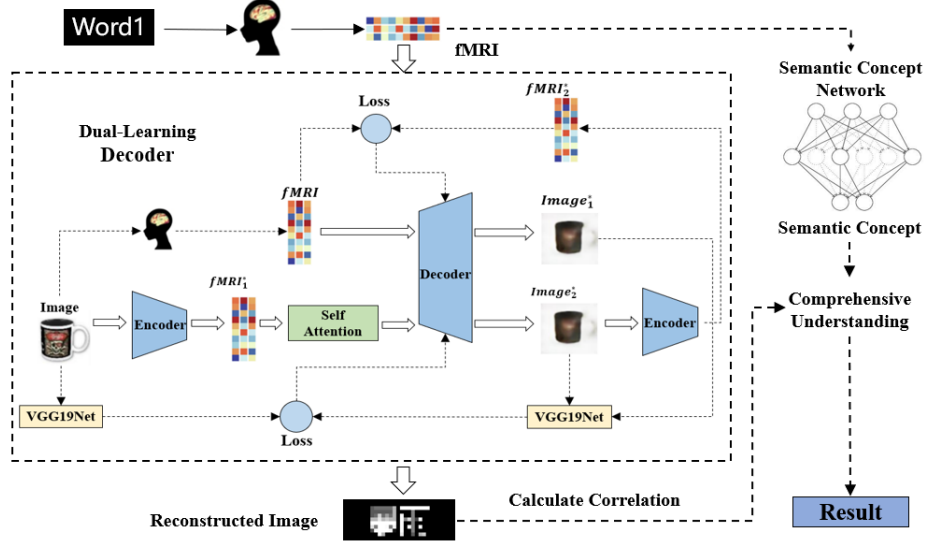


Fig. 3. The framework of Brain Signal Semantic Understanding.

4.1 Improved Dual-Learning Model

We optimize the bidirectional learning model in the algorithm model, and defines a total of four loss functions to optimize the performance of decoding training.

First, the L^E loss function between the predicted $fMRI_1^*$ data encoded by the real Image and the real brain signal fMRI data to optimize the Decoder; the second is the real brain signal. The L^D loss function between the predicted $Image_2^*$ decoded by fMRI and the real Image optimizes the Decoder. The labeled (Image, fMRI) training set is needed in the calculation and training process of these two loss functions.

Secondly, the Decoder is optimized by applying the L^{DE} loss function between the real Image and the predicted image $Image_1^*$ obtained by encoding and decoding it; Applying the real brain signal fMRI data and decoding and encoding it the L^{DE} loss function between $fMRI_2^*$ data optimizes the Decoder. This method is the Dual-learning.

And we also introduced the self-attention mechanism module in the transformer to the model for further screening of voxels to enhance effective brain signals.

Encoder and decoder are trained as follows:

We first perform supervised training of the encoder to allow it to converge in the first stage, and then provide guidance for the unstable decoding task, thereby performing encoder supervised training:

Let $\hat{f} = E(i)$ denote the fMRI response encoded by the Encoder from image i , we define the functional fMRI loss as the mean squared error and cosine near-convex combination relative to the real fMRI data f , the loss is defined as:

$$L_f(\hat{f}, f) = \rho \|\hat{f} - f\|_2 - (1 - \rho) \cos(\angle(\hat{f}, f)) \quad (3)$$

ρ is the empirical hyperparameter, and this loss is used to train the Encoder.

After the encoder training is completed, the encoder is fixed for decoder training. The training loss of the decoder in this paper includes four main losses:

$$L^{Model} = L^D + L^E + L^{ED} + L^{DE} \quad (4)$$

L^D : Decoder supervised training Given a (fMRI, Image) training pair (f, i) , a supervised loss L^D is applied to the decoded image, $\hat{i} = D(f)$, and the image reconstruction target L_i is defined by:

$$L^D = L_i(\hat{i}, i) \quad (5)$$

L_i includes the loss of image color loss L_{Color} and its features $L_{features}$. We use $\omega(i)$ to express the features of image i , and choose ω as a pretrained feature extractor, the image loss \hat{i} of the reconstructed image is:

$$L_i(\hat{i}, i) = L_{Color}(\hat{i}, i) + L_{features}(\hat{s}, s) + M(\hat{s}) \quad (6)$$

$$L_{Color}(\hat{i}, i) \propto \|\hat{i} - i\|_1, L_{features}(\hat{i}, i) \propto \|\omega(\hat{i}) - \omega(i)\|_2, M(\hat{s}) \propto SSIM \quad (7)$$

L^E : Encoder supervised training as shown, given a (Image, fMRI) training pair (f, i) , apply a supervised loss L^E , $\hat{f} = E(i)$ to the encoded fMRI, and pass the fMRI Rebuild target L_f definition:

$$L^E = L_f(\hat{f}, f) \quad (8)$$

L^{ED} : We used about 10,000 additional semantic images from the 120-concept data of the semantic image database built by the team, allowing to learn common high-level feature representations for various new concepts. In order to meet the image training without fMRI brain signal, we map the image to itself via the Encoder-Decoder transformation:

$$i \rightarrow \hat{i}_{ED} = D(E(i)) \quad (9)$$

On an unlabeled image i , the unsupervised loss L^{ED} of the loss is:

$$L^{ED} = L_i(\hat{i}_{ED}, i) \quad (10)$$

L^{DE} : We select a large number of unlabeled fMRI images in the experimental and public datasets for training, so that the network can better decode the fMRI features, we pass the Decoder-Encoder The transformation maps the fMRI response to itself:

$$f \rightarrow \hat{f}_{DE} = E(D(f)) \quad (11)$$

On an unlabeled fMRI f , the unsupervised loss L^{DE} of the loss is:

$$L^{DE} = L_f(\hat{f}_{DE}, f) \quad (12)$$

It is worth noting that the fMRI signals we use here are from the acquired dataset (which is in line as we have never used nor know their corresponding test images). This allows the decoder to generalize more to fMRI statistics. After training the decoder with these four losses $L^D + L^E + L^{ED} + L^{DE}$, we use it for decoding reconstruction of fMRI data.

The algorithm of improved Dual-learning Network is shown in Algorithm.2.

Algorithm 2 Dual-learning model

Require:

There are three groups of data needed in the training: paired(Image, fMRI) training pair $(f, i) = \{(f_1, i_1), (f_2, i_2), \dots, (f_n, i_n)\}$, a large number of semantic concept images $i^* = \{i_1^*, i_2^*, \dots, i_m^*\}$ without corresponding brain signal data selected from our semantic concept data set, and fMRI brain signals $f^* = \{f_1^*, f_2^*, \dots, f_v^*\}$, without corresponding image data collected from our experiments and a large number of public data sets $L_f(\hat{f}, f)$ Is fMRI loss function, $L_i(\hat{i}, i)$ Is the image loss function, $D(F, I)$ stands for $flag_{f \rightarrow i}$, $E(F, I)$ stands for $flag_{i \rightarrow f}$, the training step size refers to the data set itself.

Ensure:

min L^{Model} ; The reconstructed image that is most similar to the real image

- 1: $L^{Model} = L^D + L^E + L^{ED} + L^{DE}$
 - 2: $L_f(\hat{f}, f) = \rho \|\hat{f} - f\|_2 - (1 - \rho) \cos(\angle(\hat{f}, f))$
 - 3: $L_i(\hat{i}, i) = L_{color}(\hat{i}, i) + L_{features}(\hat{s}, s) + M(\hat{s})$
 - 4: **for** $a = 1$ to n **do**
 - 5: $L^D = L_f(\hat{f}, f)$, **update** $D(F, I)$
 - 6: **end for**
 - 7: **fixed** $D(F, I)$
 - 8: **for** $s = 1$ to $n + m + v$ **do**
 - 9: **for** $x = 1$ to n **do**
 - 10: $L^D = L_i(\hat{i}, i)$, **update** $E(F, I)$
 - 11: $L^E = L_f(\hat{f}, f)$, **update** $E(F, I)$
 - 12: **end for**
 - 13: **for** $y = 1$ to m **do**
 - 14: $i \rightarrow \hat{i}_{ED} = D(E(i))$
 - 15: $L^{ED} = L_i(\hat{i}_{ED}, i)$, **update** $E(F, I)$
 - 16: **end for**
 - 17: **for** $z = 1$ to v **do**
 - 18: $f \rightarrow \hat{f}_{DE} = E(D(f))$
 - 19: $L^{DE} = L_f(\hat{f}_{DE}, f)$, **update** $E(F, I)$
 - 20: **end for**
 - 21: **end for**
-

4.2 Brain Semantic Understanding Model

The image reconstruction model based on semantic concepts constructed in this paper is mainly composed of three parts: bidirectional learning reconstruction network, semantic concept cognitive network, and image semantic information discrimination network. The discriminative result of ideal semantic information cognition.

In an image semantic discrimination process, first, the brain signal of the visual brain area generated by the semantic information stimulation is input into the Dual-learning reconstruction network we constructed. The decoder trained by multiple loss functions can be obtained from the brain signal. The semantic image corresponding to the brain signal was reconstructed from the fMRI data, and its similarity with all stimulus sources was calculated.

Secondly, the brain signal of the semantic brain area generated under the stimulation of the same semantic information is input to the semantic concept cognitive network, and the semantic concept cognitive network can make a conceptual judgment on the fMRI brain signal we input, and the semantic concept category corresponding to the output brain signal.

Finally, the image semantic information discrimination network can combine the results obtained in the first two steps for further judgment, and analyze the similarity between the restored and reconstructed image and the semantic image of the determined semantic category. One is the result we need to output - the semantic information corresponding to the fMRI brain signal.

5 Experiments and Results

In this paper, a task-state fMRI experiment was designed and completed by our team. The visual stimulus images are 18 Chinese semantic text images. At the beginning of the experiment, a preparatory phase was performed first, with a blank space of 24 seconds, so as to stabilize the state of the instrument and relax the mental state of the subjects. After the preparatory phase is over, the subject of the experiment begins. In 18-s cycles, a 9-s blank was presented in each cycle, followed by a 9-s visual stimulus image before the next cycle was started. The subjects' fMRI brain signals were continuously collected throughout the experiment. Each image is presented 4 times in random order. We chose some of visual stimulus images shown in Fig. 4.



Fig. 4. Some of Visual stimulus images.

The experiment was completed in the First Affiliated Hospital of Xi'an Jiaotong University using a 3.0T Ge MR scanner. Six subjects, aged between 20 and 25, were recruited in the experiment. Three male and Three women read and signed the information sheet of magnetic resonance examination before the experiment. All subjects have normal vision or corrected vision, no history of mental illness or neurological disease, no psychological disease, no long-term medication, right-handed. Before the start of the experiment, the subjects were shown and told the content of the experiment and the tasks that the subjects needed to complete. Before the start of the experiment, the subjects were told to lie on their backs, breathe calmly, keep their heads still, stay awake, and focus on the central fixation point during the experiment.

The data collection process is as follows: First, collect T1 structural image data with the following parameters. Repetition Time (TR): 2250ms, Echo Time (TE): 3.06ms, Flip Angle: 90°, Field of view (FOV): 214mm×214mm, Voxel size: 1mm×1mm×1mm. After the acquisition of T1 structural image data, single-image stimulation experiment

and double-image stimulation experiment were carried out in turn to collect T2 functional image data. The parameters are as follows. Repetition Time (TR): 3000ms, Echo Time (TE): 30ms, Flip angle: 90° , Field of view: 214mm \times 214mm, Voxel Size: 3mm \times 3mm \times 3mm. The screen resolution is 800*800 and the refresh rate is 24Hz.

Finally, the experiment obtained 6 sets of fMRI data for 6 subjects, each set of data is a four-dimensional matrix with a size of 61*73*61*200, of which the first three dimensions are the size of the 3D whole-brain fMRI image, and 200 is the time point number.

The data preprocessing of this experiment was done by DPABI software.

The fMRI brain signals of four groups of subjects with complete experimental results and less noise in the subject data were selected for image reconstruction, and compared with the results of the Hcorr-Net [22] and unimproved bidirectional learning model previously proposed by our team. The correct rate of pixel-based image classification is shown in the Table.1.

Table 1. Results of image reconstruction.

Subject	Hcorr-Net	Dual-learning	Ours
Sub1	60.12%	62.13%	67.63%
Sub2	57.56%	58.22%	65.11%
Sub3	59.94%	59.75%	64.36%
Sub4	55.23%	60.09%	70.75%

The histogram of the correct rate of image reconstruction can better show the difference between the three, as shown in the Fig.5.

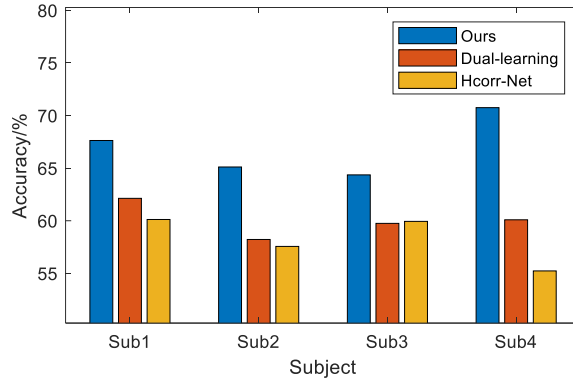


Fig. 5. Results of image reconstruction.

The fMRI brain signals of six groups of subjects were used to judge the correctness of semantic concepts, and a total of 18 groups were carried out. The correct rate of semantic concept understanding is shown in the Table.2.

Table 2. Results of semantic concept understanding.

Subject	Sub1	Sub2	Sub3	Sub4	Sub5	Sub6
Correct rate_1	94.44%	88.88%	100%	94.44%	83.33%	88.88%
Correct rate_2	100%	94.44%	94.44%	83.33%	88.88%	77.77%
Correct rate_3	83.33%	88.88%	72.22%	100%	83.33%	100%
Correct rate_4	88.88%	100%	88.88%	94.44%	77.77%	94.44%

We can clearly see from the results that compared with the previous work, the model we constructed has significantly improved in image reconstruction, and the indicators of image restoration have improved in all aspects. Moreover, our model accomplishes what other work cannot do—understands and classifies semantic concepts, and achieves high accuracy in experimental results across multiple subjects.

6 Conclusion

In this paper, our team proposes a Semantic Concept Cognitive Network (S-ConceptNet), which measures the relationship between images and brain activity data through ranking loss in multi-category learning through neural networks. The semantic concept analogy corresponding to the fMRI brain signal of the input semantic brain area can be accurately judged. Our S-ConceptNet model mainly addresses three problems: how to obtain a cognitive model of brain semantic concepts through experiments, how to fit the process of semantic concept understanding, and how to use this cognitive mechanism to guide the work of coding modeling. The results of collecting experimental data and performing image reconstruction show that S-ConceptNet effectively simulates the process of brain semantic concept understanding and reveals the cognitive mechanism related to semantic tasks. In comparative experiments, the actual effects of our proposed model are compared, and It is found that the image encoding and decoding model applying S-ConceptNet is better than the previously proposed model on the image restoration task, and can accurately output the semantic information corresponding to the brain signal, which proves that the S-ConceptNet is effective in the image restoration task. Performance improvements have been made in the corresponding areas of work. The work of this paper helps to explain the working mechanism of human semantic cognition process, and reveals the relationship model between visual information and language semantic concepts in the research of brain-like artificial intelligence. Several models proposed in this paper still have room for improvement. In future work, fMRI data sets from other related experiments can be added, and the fMRI data used can be screened for correlation or causality for further improvement. and generalize the S-ConceptNet model for further research discovery.

References

1. Gazzaniga M S, Ivry R B, Mangun G.: Cognitive Neuroscience; The biology of the mind. 3rd edn. Norton, New York (2014).

2. Robbins S.: A misdirected principle with a catch: explicability for AI. *Minds and Machines* 29(4), 495–514 (2019).
3. Zheng N, Liu Z, Ren P, Ma Y, Chen S, et al.: Hybrid-augmented intelligence: collaboration and cognition. *Journal of Zhejiang University Science C* 18(2), 153–179 (2017).
4. Logothetis N K, Pauls J, Augath M, et al.: Neurophysiological investigation of the basis of the fMRI signal. *Nature* 412(6843), 150–157 (2001).
5. Fujiwara Y, Miyawaki Y, Kamitani Y.: Modular encoding and decoding models derived from Bayesian canonical correlation analysis. *Neural computation* 25(4), 979–1005 (2013).
6. Miyawaki Y, Uchida H, Yamashita O, et al.: Visual image reconstruction from human brain activity: A modular decoding approach. *Journal of Physics Conference Series* 197(1), 012021(2009).
7. Shen G, Dwivedi K, Majima K, et al.: End-to-end deep image reconstruction from human brain activity. *Frontiers in computational neuroscience*. 13:21 (2019).
8. Horikawa T, Kamitani Y.: Generic decoding of seen and imagined objects using hierarchical visual features. *Nature communications* 8(1), 1–15 (2017).
9. Rakhimberdina Z, Jodelet Q, Liu X, et al.: Natural Image Reconstruction from fMRI using Deep Learning: A Survey. In: *Frontiers in Neuroscience*, pp. 1730. (2021).
10. Mao Z, Su Y, Xu G, et al.: Spatio-temporal deep learning method for adhd fmri classification. *Information Sciences* 499, 1-11 (2019).
11. Qureshi M N I, Oh J, Lee B.: 3d-cnn based discrimination of schizophrenia using resting-state fmri. *Artificial intelligence in medicine* 98, 10–17(2019)
12. Lin Y, Li J, Wang H.: Dcnn-gan: Reconstructing realistic image from fmri. In: *16th International Conference on Machine Vision Applications (MVA)*, pp. 1-6. IEEE (2019).
13. Noonan K A, Jefferies E, Visser M, et al.: Going beyond inferior prefrontal involvement in semantic control: evidence for the additional contribution of dorsal angular gyrus and posterior middle temporal cortex. *Journal of cognitive neuroscience*25(11),1824-1850(2013).
14. Binder J R, Desai R H, Graves W W, et al.: Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral cortex*19(12), 2767-2796(2009).
15. Jefferies E.: The neural basis of semantic cognition: converging evidence from neuropsychology, neuroimaging and TMS. *Cortex*49(3), 611-625(2013).
16. Grodniewicz J P.: The process of linguistic understanding. *Synthese*,198(12): 11463-11481.(2021).
17. Zhang Q, Wang H, Luo C, et al.: The neural basis of semantic cognition in Mandarin Chinese: A combined fMRI and TMS study. *Human brain mapping*, 5412-5423 (2019).
18. Timothy E. J. Behrens, H. Johansen-Berg, et al.: Non-invasive mapping of connections between human thalamus and cortex using diffusion imaging. *Nature Neuroscience* 6(7), 750–757 (2003).
19. Du C, Li J, Huang L, et al.: Brain encoding and decoding in fMRI with bidirectional deep generative models. *Engineering*5(5), 948-953(2019).
20. Wang H, Huang L, Du C, et al.: Neural Encoding for Human Visual Cortex with Deep Neural Networks Learning “What” and “Where”. In: *IEEE Transactions on Cognitive and Developmental Systems*, pp. 827-840 (2020).
21. Beliy, Roman, et al.: "From voxels to pixels and back: Self-supervision in natural-image reconstruction from fMRI." In: *Advances in Neural Information Processing Systems vol:32* (2019).
22. Yu S, Zheng N, Wu H, et al.: Exploring Brain Effective Connectivity in Visual Perception Using a Hierarchical Cor-relation Network. In: *Proceedings of the IFIP International Conference on Artificial Intelligence Applications and Innovations 2019*. pp.223-235 (2019).