How Does the Brain See the World – Decoding Visual Stimuli Using fMRI

Junzhu Ma

10 Mar. 2022

Abstract

The visual system is one of the most complicated sensory systems whose encoding of our day-to-day visual stimuli is still a mystery. At the start of the visual processing stream, visual information enters our retina which is later relayed to the thalamus before they finally reach visual cortices. In this project, by using data from the Kay experiment [9], I explored how the different brain areas involved in visual processing encodes different types of visual stimuli, and whether this process shares similarities with principal component analysis (PCA). By correlating brain activities with the images' projections on principal components (PCs) extracted from the Kay image dataset, I found V1 voxel activities have the highest average correlation with the images' projections on the first 5 PCs. At the same time, I also found that LatOcc process 'entity' images most similarly to PCA among all brain regions.

Introduction

Imagine you are appreciating the work of Picasso in an art gallery; or traveling in a foreign city trying to find your way to a hotel; or accidentally bumping into an old friend recalling what he looked like in the past. Such day-to-day events require complex cognitive functions where visual inputs are essential. These inputs are incorporated in the brain to assist other brain areas and their corresponding functions such as the hippocampus managing memories, the frontal lobe making decisions, and the limbic system controlling emotions. Hence, understanding the visual pathways can not only help the investigation of the visual system of the brain but can also facilitate studies of other functions of the brain including memories, spatial navigation, multi-sensory integration, etc., and these functions' relationships with the visual pathway, such as what visual input does the hippocampus receive from the visual cortex. To facilitate these cognitive functions, distinct visual signals are segregated and processed by different means.

The retina is the first hub of the visual pathway. Here, different visual information is processed and segregated. There are two mechanisms that the retina adopted to differentiate visual signals —via the same population of neurons responding differentially to different visual input, or via different populations of neurons responding specifically to certain sets of visual stimuli. For the former, there are some examples. Firstly, to detect the presence of visual stimuli, photoreceptors adopt a mechanism called the dark current. This enables them to be not activated when light is present and activated when light is not present. In this way, the information of whether light is present in a particular position of the visual field can be transmitted to the downstream photoreceptors. More complex visual signals can also be encoded at the retina level through the same principle. The direction-specific ganglion cell is another example. These neurons fire more when visual stimuli in their receptive field are moving in their preferred direction and

vice versa. In an experiment (Nelson R et al. 2001), when visual stimuli moving in four different directions (up, down, left, and right) were presented as stimuli, activities of the recorded ganglion cells only changed significantly when the stimulus is moving in the vertical direction [1].

Ganglion cells can not only identify the direction of the visual signal's movement, but also the size and shape of light in the visual signal. In an experiment, four types of visual signals are provided to a retina – no light, a small dot of light, a big dot of light, and cyclic light. The researcher recorded the activities of two types of retinal ganglion cells. For the first type of ganglion cell, it has the highest firing rate when the cyclic light is provided and the lowest firing rate when the small dot of light is provided. On the contrary, the other type of ganglion cell has the highest firing rate when the small dot of light is provided and the lowest firing rate when the cyclic light is provided. This result indicates that in this way, the size and shape of light in the visual signal can be identified in this early stage of the visual pathway. To sum up, by distinguishing visual information through different firing rates of some cells in the retina, retinal processing maximizes the amount of visual information that can be encoded in each individual unit.

Besides, the retina can also differentiate different visual signals by having them processed by different populations of cells, each in charge of encoding a specific but distinct feature of the input stimuli. Firstly, rod and cone cells process visual signals with different spatial resolutions respectively. This is achieved through the concept of convergence. When passing down visual information, many rod cells are connected to one rod bipolar cell while only one cone cell is connected to one cone bipolar cell, which means that one rod bipolar cell will receive visual information from many rod cells but cannot identify where the information is coming from. This results in lower spatial resolution of rod bipolar cells, processing a much lower spatial resolution compared to their cone counterparts. The same principle can be applied in terms of color vision.

For example, there are three types of cone cells -- blue, green, and red cone cells, each expressing different opsins sensitive to a specific wavelength of light. This color information received by photoreceptors will be sent to different color opponent bipolar cells, each responsible for detecting the boundary of a specific set of color patches. Through the mechanisms of differentiating visual signals based on the same neurons having different activity profiles and different neurons responding differently to the same signal, the retina is the first step of processing visual information. After visual information is separated in the retina by neuron activities and functions, the segregation will be maintained and sent to the visual cortex for further processing.

Visual information processing in visual cortices follows the same principle as it is in the retina. However, at the same time, signals are processed on the brain region level rather than the cellular level. Here I summarized the commonalities under three principles. Firstly, to differentiate different visual signals, the same brain region processes different visual signals differently. For example, though simple cells and complex cells can both be found in the primary visual cortex, simple cells filter orientation and spatial frequency of visual signals while complex cells filter orientation and movement directions of visual signals. Only if the input visual signal satisfies the requirements of either simple or complex cells, the two types of cells will reach their optimum firing rates respectively. It is hypothesized that the formation of the receptive field of simple cells is the result of overlapping retinal ganglion cell receptive fields, and complex cells' receptive field formation results from that of simple cells. This may be an example of hierarchical processing in visual cortices.

Secondly, different brain pathways process different types of visual information. This processing is hierarchical. Generally, for areas more in front of the processing stream of the visual pathway, information is represented less specifically, usually only including size and direction

information. Meanwhile, more towards the end of the processing stream information is represented with more details, such information, for example, includes texture, boundaries, movement, etc. On top of this hierarchical representation of visual stimuli, each specific visual area has its corresponding specific role. For instance, the object-selective neurons in the temporal cortex. These neurons can only be activated when specific shapes or objects are present. For example, if the visual signal includes a person's face, the object-selective neurons that only react to faces will be activated. The more identical the shape is to face, the more active the neurons will be [2]. This same mechanism also persists at the brain pathway level. For example, the "what" and "where" pathway also processes different types of visual information to differentiate visual signals. The "what" pathway, located in the ventral stream, is mainly in charge of object recognition and memory while the "where" pathway, located in the dorsal stream, is in charge of localization and programming of actions.

To ensure the efficiency of visual processing in the brain cortices based on the preprocessed retinal information, the segregation of retinal output needs to be maintained. This is achieved in the lateral geniculate nucleus (LGN), a thalamus region in charge of pre-filtering retinal output before cortical processing. In the LGN, different visual information is kept separate in two pathways -- the parvocellular pathway and the magnocellular pathway. The parvocellular pathway receives sustained but slow, low contrast but high spatial resolution signals, and the magnocellular pathway vice versa. On top of the segregation based on biophysical properties of the visual stimuli, positional information on the retina is also kept segregated. The primary visual cortex exhibits a topographic representation to the retina, meaning that each point of the visual cortex field can find its correspondence in the retina. Forvia, the focus point of vision, has the highest density of photoreceptors and includes most visual information, so it corresponds to most areas on the cortex.

Correspondingly, there are fewer photoreceptors and less information on the outer ring of vision, so this area corresponds to less area on the cortex. Therefore, because of the segregation mechanism in visual cortices, integration of different types of visual stimuli no longer relies on individual neuron responses but preprocessed information streams.

How can we study such a complex visual system? There are various techniques that have been developed to answer this question. In animal-based neuroscience research, scientists adopt techniques including electrophysiology, microdialysis, and two-photon imaging using which they have established the basic knowledge of visual systems. On the other hand, if the research is performed on human subjects, different techniques will be adopted. The most popular one out of all is functional magnetic resonance imaging (fMRI). This technology reveals brain-wide activities through the measurement of blood-oxygen-level-dependent signals, a parameter reflecting the concentration of oxygenated hemoglobin in different areas of the brain through their polarity profile. Compared to animal model-based technologies, fMRI provides insight into brain-wide activities and can be applied on subjects with minimal damages. For example, in 1994, the topography of the human primary visual cortex was measured by fMRI [3]. In addition, in 2004, the maximal degree of shared neural processing in visual mental imagery and visual perception were accessed and investigated through using fMRI [4]. This technology can detect a wide recording of the brain, which is essential for the understanding of encoding the complex visual world.

High spatial resolution, which is essential for investigating the encoding of various visual stimuli by specific regions or pathways of the brain, can also be best achieved by fMRI among other technologies such as electroencephalogram (EEG) and positron emission tomography (PET). EEG can detect the electrical charges caused by activities in the brain through electrodes attached

to the sculpt. However, despite direct electrical signals being sensed, this information is collected without traces of their origin, giving the absence of spatial resolution in EEG. On the other hand, by injecting the patients with a radioactive substance, PET reveals brain activities through radioactive decay of the injected compounds which is impacted by brain activities. However, it has a relatively low spatial resolution compared to fMRI. The resolution of PET is about 5 cubic millimeters while fMRI is less than 3 cubic millimeters [5][6]. However, fMRI is not the perfect technology for revealing brain activities. For example, it has a relatively low temporal resolution. FMRI has a temporal resolution of about 3 seconds while EEG has a temporal resolution of 1-5 milliseconds [7][8]. Therefore, fMRI is perfect for the study of sensory representation rather than behavior initiation. Since I want to investigate how different visual stimuli are represented in various brain regions in humans, fMRI is the best-suited technology. Therefore, in our project, we will be analyzing an fMRI dataset that consists of brain-wide voxel activities in response to various visual stimuli. I hope through this project, I can gain a deeper understanding of the visual pathway understanding which is crucial to both medical and translational applications and day-to-day life technology.

Hypothesis

Human brain processes natural visual stimuli similarly with principal component analysis (PCA) does, and different brain regions process different types of visual stimuli in distinct ways.

Procedure

1. The Kay dataset [9] is a published fMRI dataset published in 2008 on Nature. In the experiment, human subjects were asked to lay in an fMRI scanner presented with grayscaled natural

images of different categories. Meanwhile, their brain activity was recorded using fMRI. In the experiment, Kay repeatedly presented the same image 13 times and presented the subject 1750 different images altogether. The current dataset consists of the average recorded response of each voxel over the 13 repeated presentations. The voxels were aligned with the human brain atlas and the fMRI signals of 7 vision-related areas – V1, V2, V3, V3A, V3B, V4, and lateral occipital cortex – are extracted.

a) To visualize how the activities of the brain differ when the subjects are presented with different images, I plotted a heatmap, where the x-axis represents the identity of individual visual stimuli, and the y-axis is the voxel number in the visual cortex.

b) To further analyze and visualize brain activities, I plotted various heatmaps after dividing the dataset in three different ways: by the brain regions, by the semantic category of visual stimuli, and by both. This results in 7 heatmaps corresponding to the activities of 7 brain regions in response to all stimuli, 8 heatmaps corresponding to the activities of the whole brain in response to 8 semantic categories of image, and 56 heatmaps corresponding to the activity of 7 brain regions each in response to the presentation of the 8 different semantic categories. To visualize the patterns in the heatmap more clearly and make them more comparable, each voxel has its amplitude normalized to the range between -1 and 1. By comparing the average, mean, and standard deviation of all heatmaps, we can speculate whether specific brain regions react similarly to all images, whether the activity of the whole brain reacts similarly to images in the same category, and whether specific brain regions react similarly to images in the same category.

2. Reorder the dataset to visualize synchronous firing and depression in response to visual stimuli in different brain regions.

a) To better visualize correlations between signals, I used hierarchical clustering to cluster voxels with similar firing patterns to all visual stimuli based on their Pearson correlation.

b) Hierarchical clustering clustered voxels that are excited or inhibited synchronously toward individual visual stimuli together. From quantifying the size of the clusters, I can find out how many voxels respond to how many stimuli similarly. If a lot of visual stimuli trigger synchronous firing in voxels, it suggests that the voxels capture more general features of visual stimuli. At the same time, the more voxels fire synchronously, the more voxels are in charge of processing the general features.

3. According to previous research, naturalistic images preserve the same hidden statistical representation revealed by PCA [10]. As all images in the Kay dataset are naturalistic images, I extracted principal components from my image dataset using principal component analysis (PCA) [11][12].

   a) In linear algebra, finding eigenvectors of a high dimensional dataset means finding axes where data points have the highest variation. By applying PCA to find eigenvectors of the visual stimulus dataset, I examined whether there are any hidden statistical patterns across naturalistic images.

   b) Meanwhile, I computed the explained variance of the visual stimuli dataset to find the optimal number of PCs that can capture the general special features of all visual stimuli. By reconstructing images using different numbers of principal components (PCs), I confirmed the choice of the optimal number of PCs.

4. Correlate PCA projection of visual stimuli and brain activity.

a) I computed the projection of individual visual stimuli on the number of principal components I selected. This yielded a 10*1750 dataset consisting of the projection of individual images on each of the 10 eigenvectors.

b) To find how similarly each brain region processes different images from using PCA, I computed the mean Pearson correlation between voxels of different brain areas and PCA projection of all visual stimuli.

c) Similar to procedure 1.2, I also investigated correlations between the projections and various subsets of my fMRI data: whole-brain activities responding to all stimuli, activities of different brain regions responding to all stimuli, whole-brain activities responding to different categories of stimulus, and activities of different brain regions responding to different categories of stimulus. The comparison of the correlation coefficient revealed whether different brain regions encode different categories of images in a similar manner as PCA.

Results

The results of this project will answer the following four research questions: do different brain regions process different aspects of visual information; do different brain regions have different degrees of contribution in encoding semantic information carried by a visual stimulus; are there any statistical properties underlying naturalistic images; and do brain regions encode visual information in a manner similar to PCA. I will divide the following result section into four parts, each answering one of the research questions.

I first plotted the normalized brain-wide activities in response to all visual stimuli. This is shown in Fig 1a, to better visualize how each brain area fire in response to all sensory stimuli, I

plotted 7 individual heatmaps after assigning the voxels to their brain region (V1, V2, V3, V3A, V3B, V4, and LatOcc, an example is shown in fig 1b. To better visualize synchronously firing voxels, I have also clustered the data using hierarchical clustering. Fig 1c is an example of a heatmap reordered based on the result of hierarchical clustering. As fig 1d shows, the heatmap of clustered V1 voxel activities exhibits clear cluster boundaries (chunks of red and blue voxels), indicating that there are some voxels of the brain region firing in a similar fashion towards a certain set of images. The width of the cluster indicates how many voxels fire in synchronous while the length of the cluster indicates how many images are eliciting the synchronous firing. The heatmap of V1, the cortical region receiving visual information in the central visual pathway, exhibited the biggest cluster, indicating that it has the most voxel synchronous firing toward the most number of images. As we go along the central visual pathway, V2, V3, V4, and later, clusters become smaller and less obvious. This result accord to the study that areas located relatively more upstream of the visual pathway stream process less specific visual information while areas more downstream are in charge of encoding more detailed visual information. Meanwhile, V3A, V3B, and LatOcc have relatively big clusters of similar brain activities though they are relatively downstream in the visual pathway. However, their clusters are smaller compared to that of V1, meaning that fewer images stimulate similar firing rates of less number of voxels in these areas compared to V1. This phenomenon may arise from the fact that these synchronous firing voxels are the ones receiving information from their upstream areas. Another possibility is that there may be other interesting features in the images causing synchronous firing in V3A, V3B, and LatOcc. To further investigate this, data also needs to be divided based on the stimulus content.

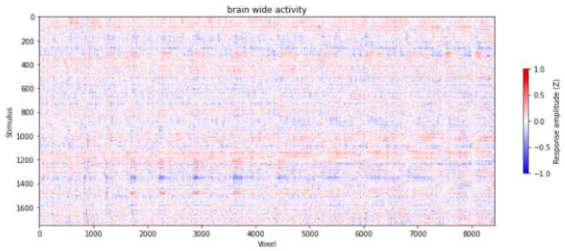

Fig 1a. heatmap of brain wide activity

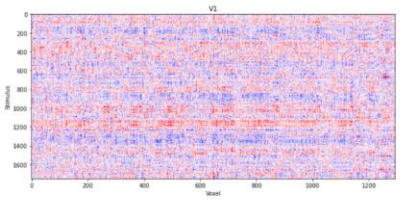

Fig 1b. heatmap of V1 brain activity

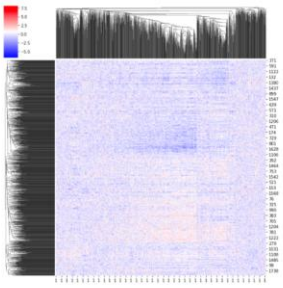

Fig 1c. cluster map of V1 activity

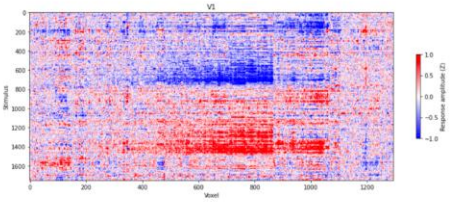

Fig 1d. heatmap of V1 brain activity after clustered

I first divided my images into general semantic categories. The Kay dataset provides labels for each image in the dataset generated by BERT [13]. Each image has a corresponding label and three levels of classification from the most general one to the least. The most general form divides all images into 8 semantic categories: animal, plant, fungi, fruit, geological formation, person, entity, and artifact. Since these labels are auto generated by a deep learning network, the labels are sometimes not accurate enough. Therefore, I checked all labels manually to make sure that they match the content of the image and changed the ones that do not match. At first, I plotted the heatmap and performed hierarchical clustering on the dataset containing brain-wide voxel activities to different categories of images. As fig 2a shows, there are many small clusters among each semantic category of stimuli but which brain region the signal comes from is unclear. Therefore, I further divided the voxel activities based on the brain regions that these voxels belong. This can help me investigate specifically how different content of images affect the activities of different brain regions. Fig 2b shows the clustered heatmaps of four example categories of images each of which categories have four heatmaps corresponding to the activities of V1, V2, V3, and

V4 to that category of image. Similar to the result of the heatmap of brain regions, heatmaps of all categories of images also exhibit decreasing cluster size down the stream of the central visual pathway. Using the 8 biggest categories to sort images is only the first step of understanding how the brain encodes visual semantic information, and it cannot best explain some specific categories in relation to brain decoding visual stimulus, because, for example, dogs and elephants all belong to the category of animal, but they are still very different. Therefore, further analysis can be done based on more detailed categories to obtain more accurate and specific results.

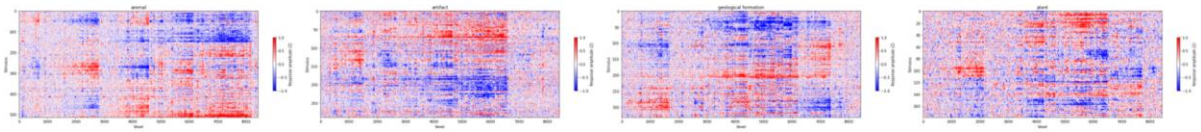

Fig 2a. heatmaps of brain-wide activities stimulated by images under the category of "animal", "artifact", "geological formation", and "plant"

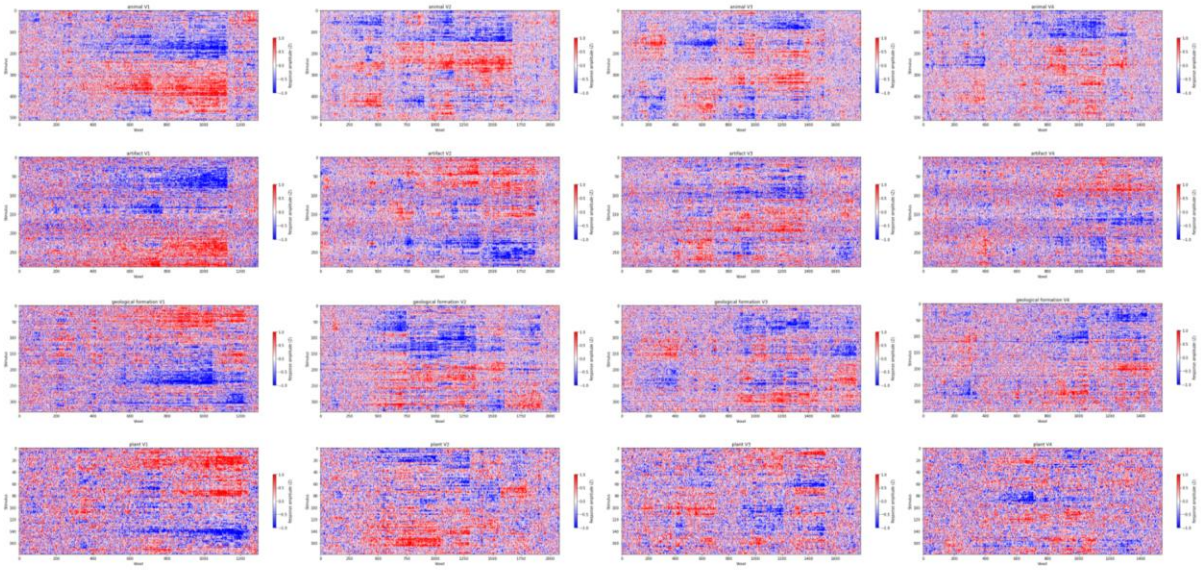

Fig 2b. heatmap of V1, V2, V3, and V4 activities stimulated by images under the category of "animal", "artifact", "geological formation", and "plant"

My next step is to explore whether there are any hidden statistical properties of the image dataset by reducing the dimensionality of the dataset. This can be captured by adopting principal component analysis (PCA). PCA is the process of computing principal components (PCs). The principal components of a dataset are eigenvectors of the input dataset. In terms of linear algebra, the first eigenvector of the dataset is the high dimensional axis where all data points have minimal

average projections, meaning that it represents the most general information representation of all data. At the same time, the later eigenvectors have a greater average projection of the data to the vector, meaning that they capture more detailed representations of the dataset. By using a few eigenvectors that have the greatest explained variance computed using PCA to represent the dataset, the dimensionality of the dataset can be reduced. To visualize the function and effect of PCA, the first 10 eigenvectors are plotted in fig 3a and averaged explained variance of the first 10 PC is plotted in fig 3b. The later the principal component is extracted the less variance of the image dataset it can explain (fig 3a), meaning that it encodes more detailed information about image features. The first 10 PCs I found from my image dataset is the same as the result of a previous research which shows the first 10 PC of all nature images [10]. To visualize how much information is preserved in the first 10 principal components, I reconstructed several images in my visual stimuli dataset. Fig 3c is one of the examples. Construction using the first 10 PCs captured the general spatial features of the image. At the same time, I also used the 300 PCs to reconstruct the same image which adds more refined details to the reconstruction without changing the general spatial feature of the image. Fig 3c and 3d are reconstructed images based on the first 10 and 300 PCs. I chose to use the first ten eigenvectors for the next part of my analysis because I am more interested in investigating the encoding of general spatial features of the images.

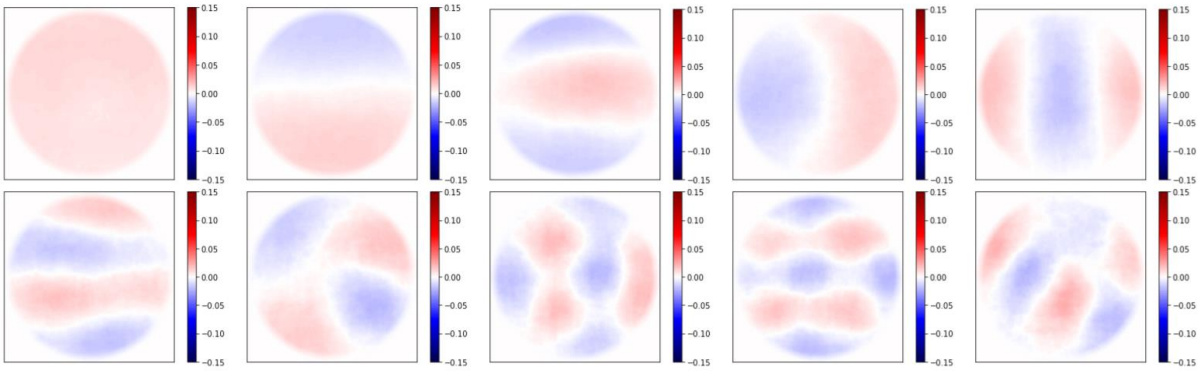

Fig 3a. the first ten eigenvectors (1st - 5th and 6th - 10th from left to right)

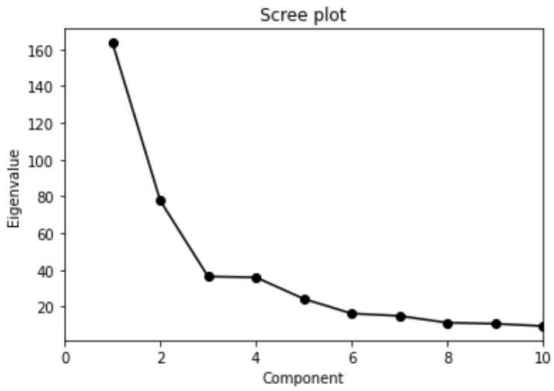

Fig 3b. the first ten explained variance (eigenvalues)

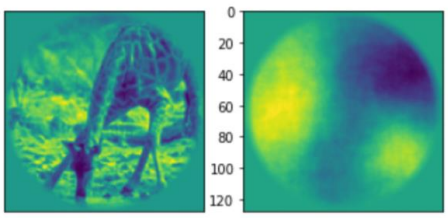

Fig 3c. reconstructed image of 10 pc

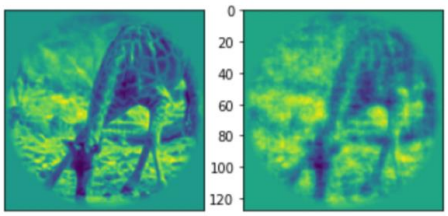

Fig 3d. reconstructed image of 300 pc

Inspired by previous research that revealed how eigenfaces can be encoded in face patches in the macaque cortex [14], I decided to compare the brain activities stimulated by naturalistic images with their projections to the dissected eigenvectors by computing the mean Pearson correlation between them. The output matrix from my customized function contains the correlation between every brain activity with each of the first 10 PCs I obtained from the image dataset. The result is shown in fig 4. Like what I did when plotting the heatmaps previously, I calculated these mean correlations by dividing the dataset in three ways: by brain regions, by image categories, and by both. My result showed that V1 has the average highest correlation with the first 5 PCs (fig 4a). At the same time, as I move to later PCs, brain regions more downstream in the visual pathway become more correlated, such as V3, which highly correlates with image projections on the 7th,

9th and 10th PC. This result supports the fact that V1 processes the most general features of images since the image projections on the first few PCs are most correlated with V1. As I move to the later PCs, they carry more detailed information of the images, and therefore, are more correlated with more downstream visual areas. Interestingly, by dividing the dataset based on semantic categories of images, I found that brain activities responding to the category 'entity' yielded the highest correlation and the category 'animal' yielded the lowest with their image projections on the first 10 PCs. This difference may arise from differences in the variability of their intra-category image information. Since brain activities corresponding to 'entity' images have the highest correlation with the image projections on all PCs, I only further divided this dataset based on the different brain regions. As a result, LatOcc has the highest average correlation with the PCs of 'entity' images (fig 4b) indicating that among all brain regions I analyzed, it has the most similar way of processing images with PCA does on 'entity' images.

PC	Brain Regions	Image Categories
1	V1>V2>V4>V3>LatOcc>V3B>V3A	entity>fruit>fungus>plant>person>gf>artifact>animal
2	LatOcc>V1>V2>V4>V3>V3B>V3A	entity>fruit>fungus>plant>gf>artifact>person>animal
3	V1>V2>LatOcc>V3>V3B>V4>V3A	entity>fruit>fungus>plant>artifact>gf>person>animal
4	V1>V2>V3>V3B>V3A>V4>LatOcc	entity>fruit>fungus>plant>person>gf>artifact>animal
5	V1>V3B>V4>V2>V3>LatOcc>V3A	entity>fruit>fungus>plant>gf>person>artifact>animal
6	LatOcc>V2>V1>V3>V3A>V4>V3B	entity>fruit>fungus>plant>gf>person>artifact>animal
7	V3>V1>V2>V3A>LatOcc>V3B>V4	entity>fruit>fungus>plant>person>artifact>gf>animal
8	V4>V3A>LatOcc>V3>V3B>V2>V1	entity>fruit>fungus>plant>gf>person>artifact>animal
9	V1>V3A>V3B>V4>V2>V3>LatOcc	entity>fruit>fungus>plant>gf>person>artifact>animal
10	V3>V1>V3B>LatOcc>V4>V2>V3A	entity>fruit>fungus>plant>person>gf>artifact>animal

Fig 4a. order of correlation of the first ten PC with brain regions and image categories

PC	Overlap ('entity')
1	V3A>V3>V2>V4>LatOcc>V1>V3B
2	LatOcc>V4>V2>V3>V3B>V3A>V1
3	LatOcc>V3>V2>V4>V1>V3B>V3A
4	V4>V3>V1>LatOcc>V2>V3B>V3A
5	LatOcc>V3A>V4>V2>V3A>V3>V1
6	V3A>V2>V3>V3B>V4>LatOcc>V1
7	LatOcc>V3>V4>V2>V1>V3B>V3A
8	LatOcc>V3A>V4>V2>V3B>V3>V1
9	V1>LatOcc>V4>V2>V3B>V3>V3A
10	V2>LatOcc>V3>V1>V4>V3B>V3A

Fig 4b. order of correlation of the first ten PC with brain regions activated by 'entity' images

Discussion

Despite yielding positive results from the project, I still found some limitations in my analysis. Firstly, some image labels are not very precise because I was unable to identify the content of them due to the level of complexity of the semantic information they carry. For example, there is an image with an old man with two children, and all of them wore cowboy hats, which makes it difficult for me to assign it into either the 'person' or the 'entity' category. As a solution, I always chose the most obvious semantic component I identified in the image to be the label. However, this can cause inaccuracies since the subject in the Kay experiment might not first recognize the same component of the image as I did. At the same time, due to the mixed semantic information in some images, the subjects could also have their brain activities encoding mixed visual information. Further analysis can investigate how these ambiguous images affect the activities of the brain by separating them to a category of their own, e.g., a woman wearing Indian ornaments can be assigned to the category of "person+artifact". Using general semantic categories to divide image stimuli is another limitation of this project. My next step is to further investigate how more detailed semantic information of the image can affect brain activities by using natural

language processing algorithms like Word2vec. Word2vec, for example, can assign a 300-dimensional vector to its word input which can make the differences between more detailed semantic information encoded in images more quantifiable. In the future, I want to investigate similarities between the semantic distances of visual stimuli with their respectively evoked brain activities.

Studying the visual pathway is the first and most important step of understanding the process of daily visual inputs. This knowledge can be applied in both medical and translational applications and day-to-day life technology. The most straightforward use of visual pathway knowledge is on treating diseases such as retinal damage and cortical blindness. Besides, some day-to-day life technologies are also based on the understanding of visual encoding. The VR industry is a great example. Virtual reality (VR) uses the ability of the brain that can turn 2D images into a stereoscopic environment to construct an immersive 3D environment. The technology cannot be implemented without understanding our visual system e.g., how binocular vision and shades contribute to depth sensing. Therefore, understanding the visual encoding process is essential to the development of disease treatments and other state-of-the-art technologies. I hope that through my project, I can also contribute to our understanding of the highly complex visual system.

Reference

1. Kolb. *The Organization of the Retina and Visual System*. Webvision, 2011.
2. Leopold, David A, and Gillian Rhodes. “A comparative view of face perception.” *Journal of comparative psychology (Washington, D.C. : 1983)* vol. 124,3 (2010): 233-51. doi:10.1037/a0019460
3. Engel, S. A., Rumelhart, D. E., Wandell, B. A., Lee, A. T., Glover, G. H., Chichilnisky, E.-J., & Shadlen, M. N. (1994). fMRI of human visual cortex. *Nature, 369*(6481), 525. https://doi.org/10.1038/369525a0
4. Ganis, Giorgio, William L. Thompson, and Stephen M. Kosslyn. "Brain areas underlying visual mental imagery and visual perception: an fMRI study." *Cognitive Brain Research* 20.2 (2004): 226-241.
5. Catana, Ciprian et al. “Simultaneous in vivo positron emission tomography and magnetic resonance imaging.” *Proceedings of the National Academy of Sciences of the United States of America* vol. 105,10 (2008): 3705-10. doi:10.1073/pnas.0711622105
6. Huber, Laurentius, et al. "Techniques for blood volume fMRI with VASO: From low-resolution mapping towards sub-millimeter layer-dependent applications." *Neuroimage* 164 (2018): 131-143.
7. Glover, Gary H. “Overview of functional magnetic resonance imaging.” *Neurosurgery clinics of North America* vol. 22,2 (2011): 133-9, vii. doi:10.1016/j.nec.2010.11.001
8. Polich, John. *Detection of Change: Event-Related Potential and fMRI Findings*. Softcover reprint of the original 1st ed. 2003, Springer, 2012.
9. Kay, K., Naselaris, T., Prenger, R. *et al.* “Identifying natural images from human brain activity.” *Nature* 452, 352–355 (2008).

10. Simoncelli, E P, and B A Olshausen. “Natural image statistics and neural representation.” *Annual review of neuroscience* vol. 24 (2001): 1193-216. doi:10.1146/annurev.neuro.24.1.1193
11. F.R.S., Karl Pearson. “LIII. On lines and planes of closest fit to systems of points in space.” *Philosophical Magazine Series 1* 2: 559-572.
12. Hotelling, H. “Analysis of a complex of statistical variables into principal components.” *Journal of Educational Psychology*, vol. 24, no. 6, 417–441.
13. Devlin, Jacob, et al. "Bert: Pre-training of deep bidirectional transformers for language understanding." *arXiv preprint arXiv:1810.04805* (2018)
14. Chang, Le, and Doris Y Tsao. “The Code for Facial Identity in the Primate Brain.” *Cell* vol. 169,6 (2017): 1013-1028.e14. doi:10.1016/j.cell.2017.05.011