University of Tehran

Campus of Technical Faculties

Faculty of Electrical and Computer Engineering

Analyzing the Primate Visual Cortex and Comparison with Convolutional Neural Networks (CNNs)

Thesis for Receiving Bachelor's Degree

In Electrical Engineering with Concentration in Control Systems

Student: Erfan Mirhaji

Student ID: 810196568

Supervisor:

Professor Mohammad-Reza A. Dehaqani

Summer of 2022

Affidavit of originality of the work

I, Erfan Mirhaji, confirm that the contents of this thesis are the result of my own efforts and the research achievements of others that have been used in this article have been referenced according to the regulations. This thesis has not been submitted before for obtaining any degree of the same or higher level. All intellectual and material rights of this work belong to Technical College of Tehran University.

Student: Erfan Mirhaji

Student's Signature:

Abstract

In the recent years, neural networks have been able to achieve a performance nearly identical to the Visual Cortex of primate brain in recognition and classification of objects [1-4]. Also, recent studies have shown that the initial layers and the final layers of a convolutional neural network are similar to the initial and final layers of the Visual Cortex of the primate brain [5-10].

Some neural networks, especially convolutional neural networks or CNN for short, simulate the known structure of the initial layers of the Visual Cortex of primate brain and iterate these layers one after another. This is repeated until the neural network is formed. The terminal and more advanced parts of the Visual Cortex of the primate brain are mostly unknown, but there is hope that the structure and function of the advanced layers of the primate brain would automatically emerge in neural networks. If so, comparing the structure of neural networks with the Visual Cortex can provide us with a shortcut to study and discover the structure and function of the advanced layers of the Visual Cortex of the primate brain.

Here, we compare the data from the Inferior Temporal Gyrus cortex (IT for short) of macaque's brain with the output of all layers of several Convolutional Neural Networks (CNNs) for identical images, and we try to verify the claim that the final layers of primate Visual Cortex and neural networks are similar.

Moreover, we measure the effect of frequency (such as whether the image is blurry or sharp) on the behavior and output of neural network layers as well as the macaque's brain. The reason for conducting these studies is to better and more clearly understand the function of the Visual Cortex and to find similarities and differences between the Visual Cortex and neural networks. This will help us, firstly, to study the advanced layers of the Visual Cortex in detail and second, design new neural networks in such a way that are as functionally similar to the Visual Cortex of the brain as possible.

The results obtained in the end are graphs of the Visual Cortex and 9 different neural networks, which show that the frequency of images greatly affects the neural networks' performance, whereas this effect on the Visual Cortex's performance is insignificant. We also concluded that the correlation of CNNs and the Visual Cortex is affected by the image frequency, and there isn't necessarily a 'similarity' between layers of CNNs and the Visual Cortex.

Keywords:

Neural Network, Convolutional Neural Network, Artificial Neural Network, Representational Similarity Analysis, Representational Dissimilarity Matrix, Inferior Temporal Gyrus cortex, Brain Visual Cortex, Image Frequency Effect on Visual Cortex, ANN, CNN, RNN, IT cortex, Visual Cortex Modeling, Primate Visual Cortex

Table of Contents

List of Abbreviations

NN	Neural Network
ANN	Artificial Neural Network
CNN	Convolutional Neural Network
RNN	Recurrent Neural Network
IT Cortex	Inferior Temporal Gyrus cortex
RSA	Representational Similarity Analysis
RDM	Representational Dissimilarity Matrix
VC	Visual Cortex

Chapter 1:

Introduction

1-1 Prologue

Our goal is to investigate and compare the function of the Visual Cortex (VC) of macaque's brain with convolutional neural networks. According to recent studies [5-8], the Visual Cortex of primates and neural networks are very similar. This means the initial layers of the VC are similar to the initial layers of neural networks, and the final layers of the VC are similar to the final layers of neural networks. Our information about the advanced and final areas of the VC is insignificant from a neurological point of view, and we do not have information about their function, but there we have detailed knowledge about the primitive and simpler parts of the VC. Comparing the function of neural networks and the VC can help us understand the function of the final layers of the VC, so that in case of similarity, we can use the information about of the structure of neural networks to estimate the function and structure of the more complex areas of the VC. Also, examining the effect of changing the frequency of images on the output of the brain as well as the output of neural networks can help us better understand the difference or similarity between the function and structure of the two, which we have discussed. In this research, we use pictures with the following frequencies: Intact, Blurry and Sharp and record the output of the VC and neural networks and analyze their performance.

1-2 A History of the Research Topic

In a recent 2014 study [8], the IT region of the VC was compared with different parts of different neural networks and according to the extracted data, it was concluded that neural networks can simulate the IT region of VC, and that the function of the IT region is comparable to the layers of the neural network. However, in a paper published in 2021 [11], a comparison was made between different areas of the human VC with different layers of several neural networks. The output of different parts of the VC of the brain have been compared with the outputs of different layers of several neural networks. The result of this research is that, contrary to the previous claims that the more advanced and final layers of VC are functionally similar to the final layers of neural networks [1,8], these two have little similarity in the final layers, although the initial layers of the VC and the initial layers of neural networks are comparable.

1-3 Description of the Research

As mentioned earlier, our goal is to compare the output of all layers of several neural networks with the output of the IT cortex of the VC of macaque's brain. For this, data was collected from the IT cortex of a macaque's brain at Prof. A. Dehaqani's Laboratory [20]. This was done by directly inserting electrodes into the macaque's brain. In each data collection session, 81 images consisting of 27 images, each in three frequencies of intact, blurry and sharp were shown to a macaque and the data was recorded. Each image was repeated 5 times during each session. That is, in each session, 81 images were shown to a macaque 5 times each, and the data of the neurons of different parts of the IT region of the VC were measured each time. This work was repeated in 169 sessions, and in each session, different parts of the IT area were examined, resulting in a total of 352 neurons' data being collected (a minimum of 1 neuron and a maximum of 6 neurons per session). Finally, to reduce the error and ensure that the activity of the neurons is not random, the average of 5 repetitions has been calculated for each of the 81 images.

In each session, the activity level of each neuron was recorded in every millisecond, from 300 milliseconds before the image was shown to the macaque until 600 milliseconds after the image was shown. Then, to calculate the activity level of each neuron, the data from the period of 70 milliseconds after displaying the image until 200 milliseconds after displaying the image has been averaged to obtain the activity of that neuron. In the end, the collected data is the activity of 352 neurons of the IT region of the VC for 81 different images, of which 27 images are intact, 27 sharpened and 27 blurry.

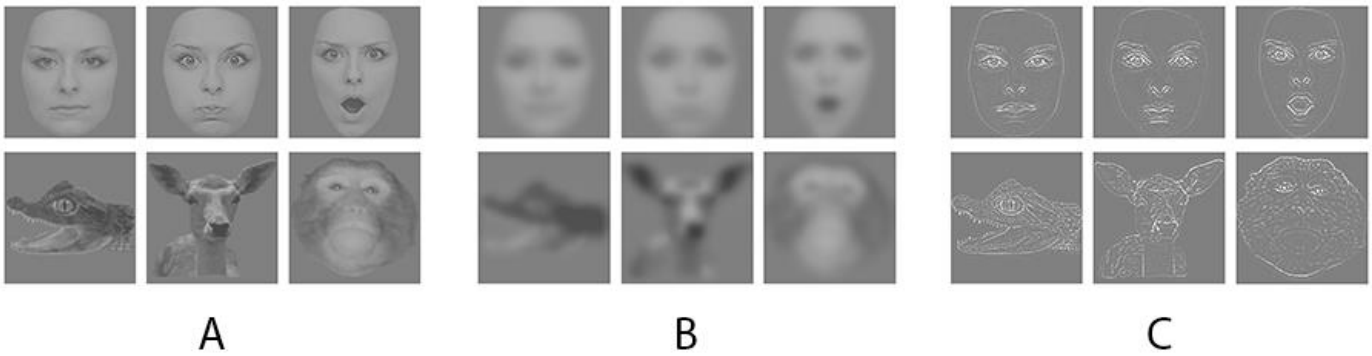

Figure 1: An example of the 27 images, each of which is displayed in 3 frequencies. A total of 81 images have been used. Fig. A intact images, Fig. B blurry images and Fig. C sharp images

Figure 2: 81 images were shown 5 times each in 169 sessions to a macaque and the activity of different parts of the IT region of the VC was recorded.

Then, for each photo, the output of each neuron is averaged so that the measurement error is reduced and the data can be reliable.

For neural networks, these 81 images are given as an input to the neural networks and the output of each layer of each neural network is recorded. Each neural network has between 25 and 500 layers, depending on its size. In the last layers, each neural network has a classifier that calculates the category to which the image belongs. All neural networks used in this research are neural networks already trained on ImageNet images and their parameters have not been changed.

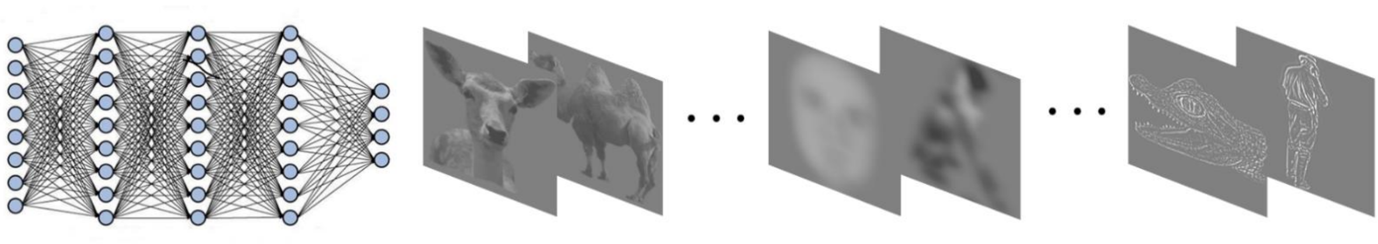

Figure 3: 81 images are given to the neural network and the output of each layer is saved.

To compare the effect of frequency on the IT region of the VC as well as neural networks, a classifier was trained, whose input is the output of the last layer before the classification layer of the neural network for images of a frequency along with their labels. And to test it, is to test the output of the same images in another frequency of the same layer along with their labels. For example, the classifier was trained with the data of blurry images and the data of sharp images was tested with it. For the brain, a similar work was done, the data related to the IT region of the VC for the images of one frequency along with their labels are given to the classifier and the test is performed by the output of the same images in another frequency from the IT region with their labels.

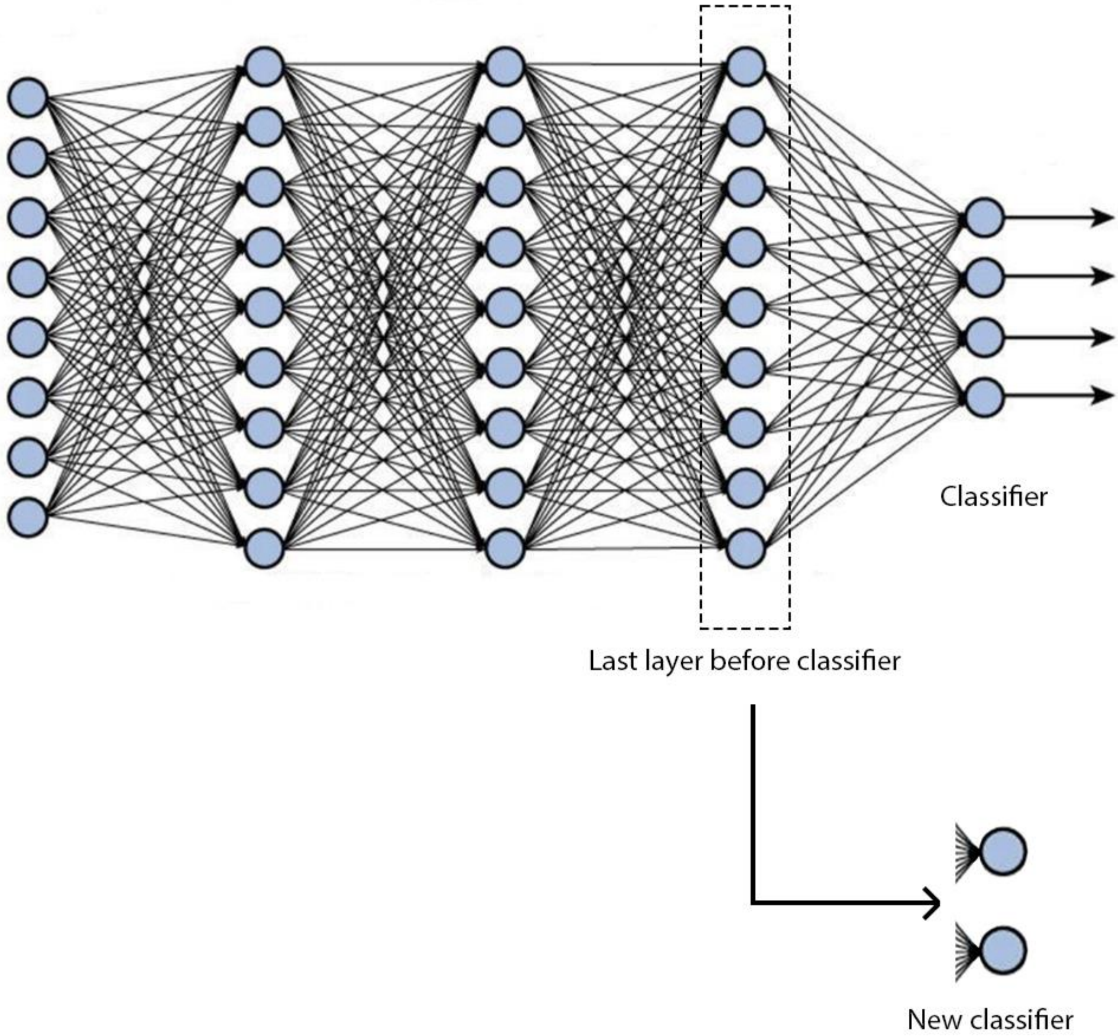

Figure 4: To analyze the effect of frequency on the neural networks, a classifier is trained, whose input is the output of the last layer before the classification layer of the neural network for images of one frequency along with their labels. For example, the classifier was trained with the data of blurry images and the data of sharp images was tested on it.

1-4 Definition of the Research Topic

As stated previously, our goal is to compare the output of the layers of several neural networks with the output of the IT region of macaque's VC. This is done by Representational Similarity Analysis (or RSA for short). This way, the data of brain 's neurons is compared with each of the layers of each of the neural networks. For this purpose, first the Representational Dissimilarity Matrix (or RDM for short) has been created for both the brain and each layer of neural networks and these two matrices have been compared with each other. The comparison has been made by calculating Pearson's Rank Correlation between the RDM matrix of brain's neurons and RDM matrices of neurons of each layer of the neural networks. The higher this correlation value is, the more similar that layer is to the brain and vice versa. This comparison has been made for all layers of the neural networks with the IT layer of VC.

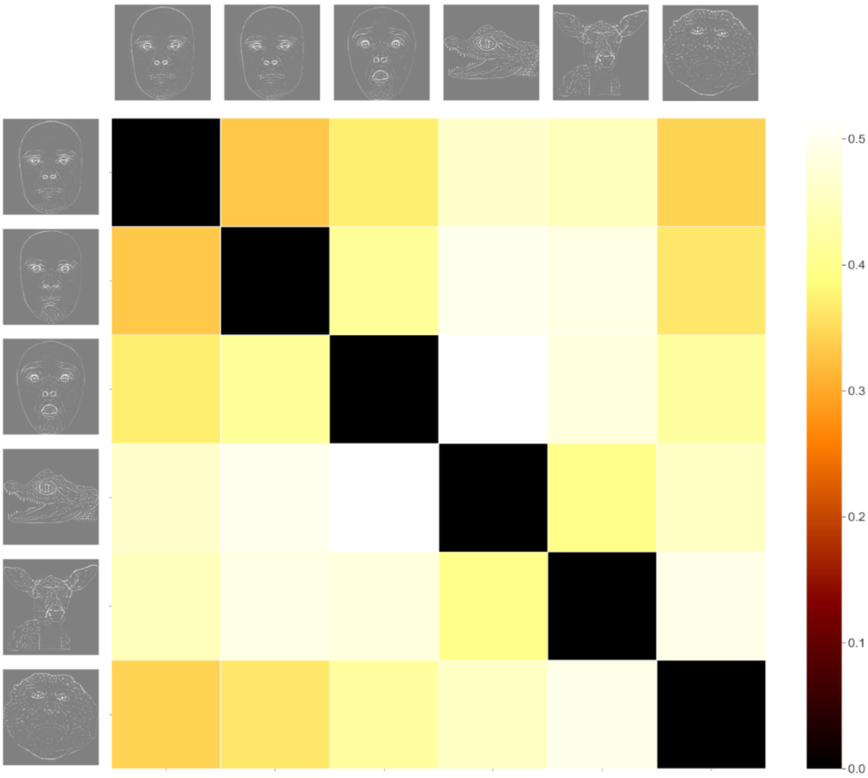

Figure 5: An example of an RDM, this RDM is for the IT region of the macaque's brain. The rows and columns of this matrix are the images that are used. Here are 6 images as examples. Each cell is the distance of the two corresponding images. The diameter of the matrix is the distance of each image from itself, which is equal to zero.

To compare the effect of frequency on the IT region of the VC and neural networks, an SVM classifier has been calculated for the brain. To train this SVM classifier, the data of the neurons for each of the intact, sharp and blurry images, along with their labels, is given. To test this SVM classifier, the data of the neurons for other frequencies along with their labels are tested on it. For neural networks, this work is done for the layer before the last classifier in the network. For training an SVM classifier, the data of the layer before the classifier of the network, for each of the intact, sharp and blurry images are trained along with their labels and to test the SVM classifier, the data of the layer before the classifier in the neural network for other frequencies along with their labels is tested.

1-5 Goals and Objectives of the Research

We seek to find the similarity of the IT region of the macaque brain with different layers of different neural networks. Also, we want to measure the effect of frequency on the functional behavior of different layers of neural networks as well as the IT region of the VC. If these two are similar, we can use neural networks to understand and analyze the structure and function of the IT region of the VC. If these two are not similar, first, we cannot use neural networks to analyze the more advanced and final layers of the brain's VC, and secondly, we should look for a design of neural networks that in its final layers, acts similar to the final layers of the VC of the brain. If a neural network simulates the function of the Visual Cortex of the brain, there should be a functional correspondence between the layers of the neural network and the VC of the brain. There should be a similarity between the two and the results that we get from examining the brain should also be seen in the neural network.

1-6 Research Procedure

To carry out this project, first the concepts of machine learning, neural networks and deep learning were learned. Then, studies were conducted to understand the brain structure of primates and the function and task of each of the layers of the Visual Cortex of the brain. Then, to work with brain's data and filter and sort and process them, data processing tools and data analysis were learned. In the next step, I studied pre-trained neural networks and worked with their tools. There was also a need to learn the methods of comparing and performing operations on neural network data as well as the extraction of the data, which was done. Finally, tools and methods of working with images, pre-processing of input data, processing of output data of neural networks and comparison of brain data with neural network data were studied and comparisons were made.

1-7 Thesis Structure

This report includes the following chapters:

Chapter 2:

The concepts and the background needed to understand the function of the Visual Cortex of the macaque brain and neural networks as well as how to compare them are discussed.

Chapter 3:

We discuss the proposed method of working with the neural network and its data, as well as working with the data of the visual area of the monkey brain and comparing the two with different methods.

Chapter 4:

Description of the results of comparisons and calculations and announcement of the results obtained from the investigations.

Chapter 5:

Summarizing and presenting suggestions for improving the performance of neural networks

1-8 Summary

In this chapter, we explained the purpose of the research, the research method and the final result of the research. The purpose of the research is to compare the data of the IT cortex of macaque's brain with the output of different layers of several neural networks and to verify the claim that the final layers of the brain and final layers of networks are comparable. Also, we measure the effect of image frequency on the behavior and output of neural network layers as well as the brain. The method of doing this comparison is through RSA analysis and correlation comparison of brain's data and neural networks. The result of the research is that the function of the IT region of the macaque's brain and neural network layers are different and also the effect of frequency on the brain's function is insignificant, but the performance of the neural networks changes.

Chapter 2:

Basic Concepts and Prerequisites

2-1 Introduction:

In this chapter, we explain the basic concepts and requirements for conducting this research and comparing the IT region of the brain with neural networks. First, we need to know the structure of macaque's brain and the function of its different regions. Then we need to examine the structure and function of different neural networks. Finally, we need to grasp the tools that allow us to compare the two.

2-2 Macaque's Brain

The Visual Cortex of the human brain and in general the brain of primates has several different layers, both physically and functionally. The initial layers have simpler functions, and the more advanced and final layers have more complex and advanced functions [12-15]. Primitive and initial regions include V1, V2, V4 etc. The terminal regions that are more advanced include LOT and IT [12,16].

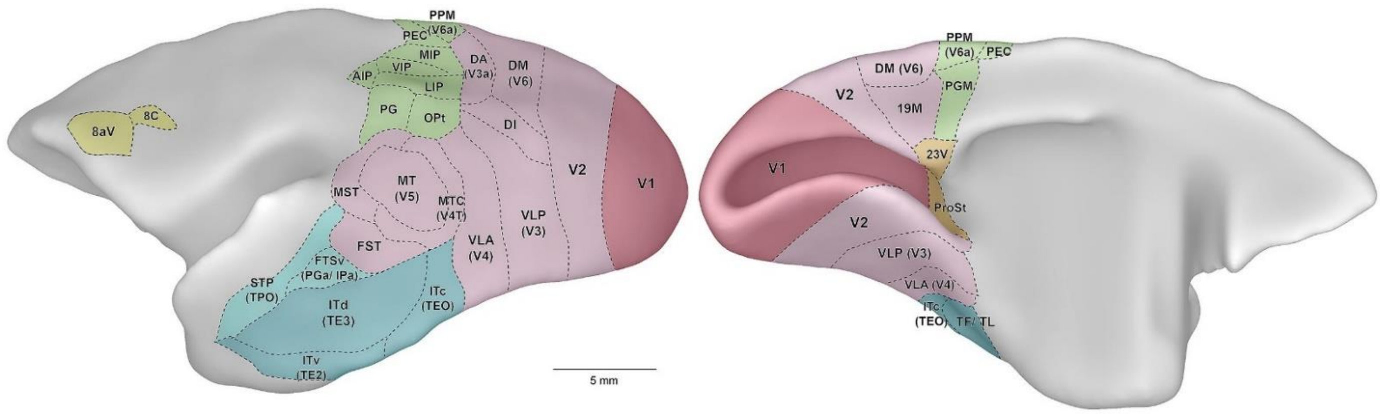

Figure 6: Macaque brain and different regions of the Visual Cortex [12]

The Primary Visual Cortex region is located in the posterior pole of the occipital lobe and is the region that has been studied the most. This area specializes in processing still and moving objects and excels at pattern recognition. This area is classified into 6 parts, which are known as V1, V2 and ... V6.

The region of the Inferior Temporal gyrus or IT for short is located under the Middle Temporal gyrus and is connected to the Inferior occipital gyrus. The IT area includes 4 parts: FFA, PPA, EBA and LOC, which are closely connected with each other together with the Hippocampus. This region is connected to V1, V2, V3 and V4 regions and processes the information received from them. The IT region processes the color and form of objects and its task is to recognize what an object is. That is, what is the image that is seen by the eyes.

2-2 Neural Networks

Neural networks, which have made great progress in recognizing and classifying images in the recent years, consist of different layers, each of which has a specific task. The primary and initial layers are designed to recognize simple information such as horizontal and vertical lines, and the later layers are designed to recognize more advanced features of the environment and the position of objects. Neural networks are of several types, such as ANN, CNN and RNN. These networks are consisted of a large number of neurons, each of which performs simple calculations and is connected to other neurons depending on the function of the network. Some networks, such as CNNs, inspired by the structure of the initial layers of the Visual Cortex of the brain, have simulated this structure and repeated it several times over. In this research, we compare the performance of CNN networks with the IT region of the brain. The networks used in this project are: AlexNet, GoogleNet, MobileNet-V2, ResNet18, ResNet50, Resnet101, VGG-16, VGG-19 and SqueezeNet1.1.

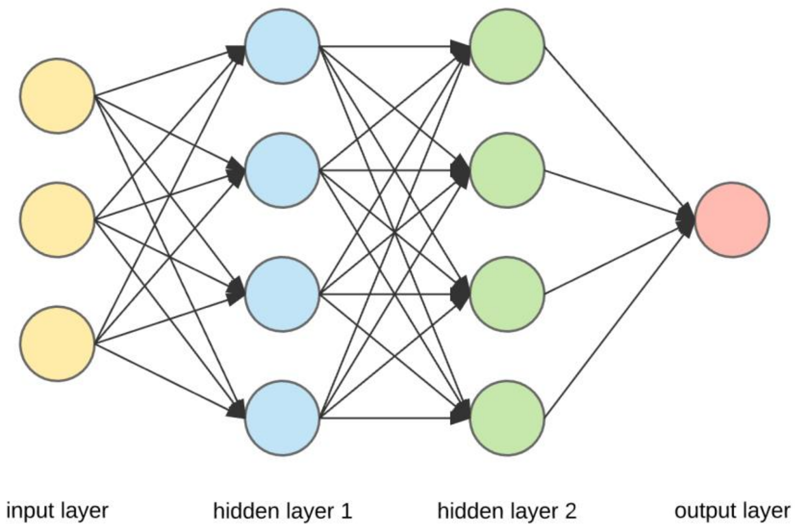

Figure 7: An ANN neural network [17]

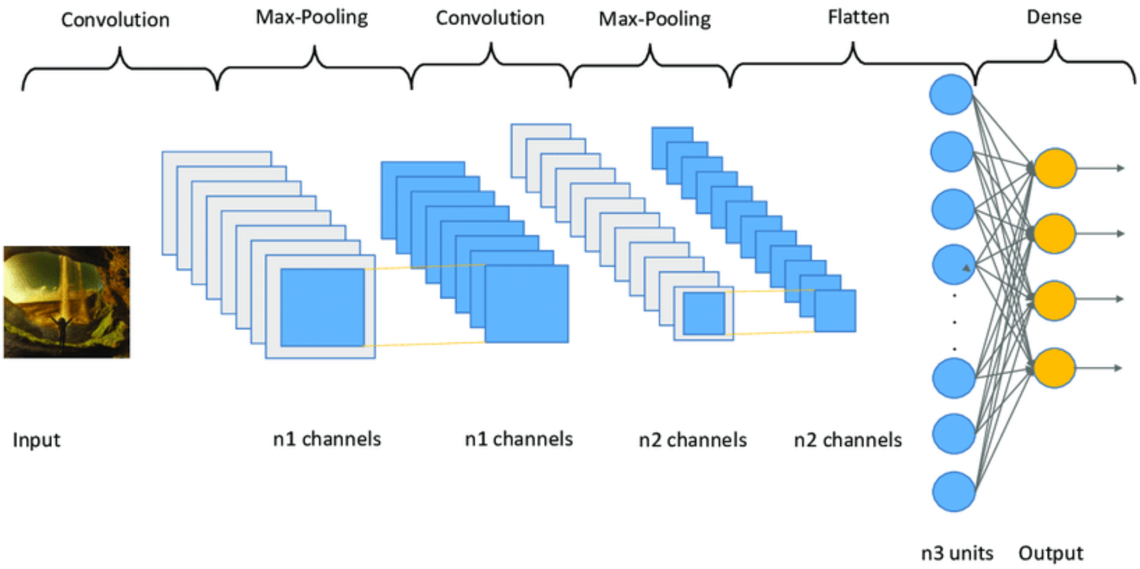

Figure 8: A CNN neural network [18]

Below we can see the output of different layers of a CNN network that was used in this research:

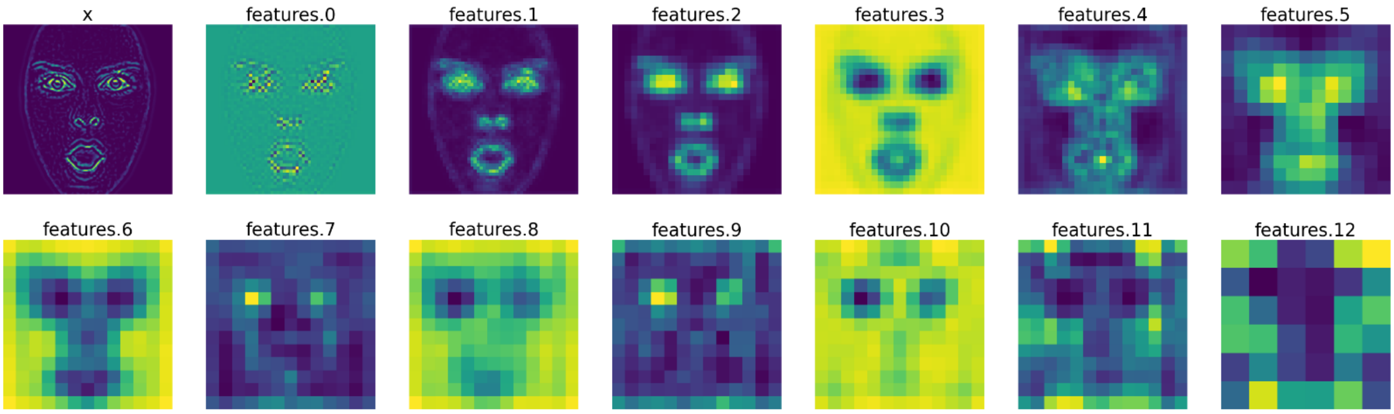

Figure 9: An example of the convolutional outputs of the AlexNet network, which is a CNN. X represents the input of the neural network, and the rest of the pictures are the output of the convolutional layers

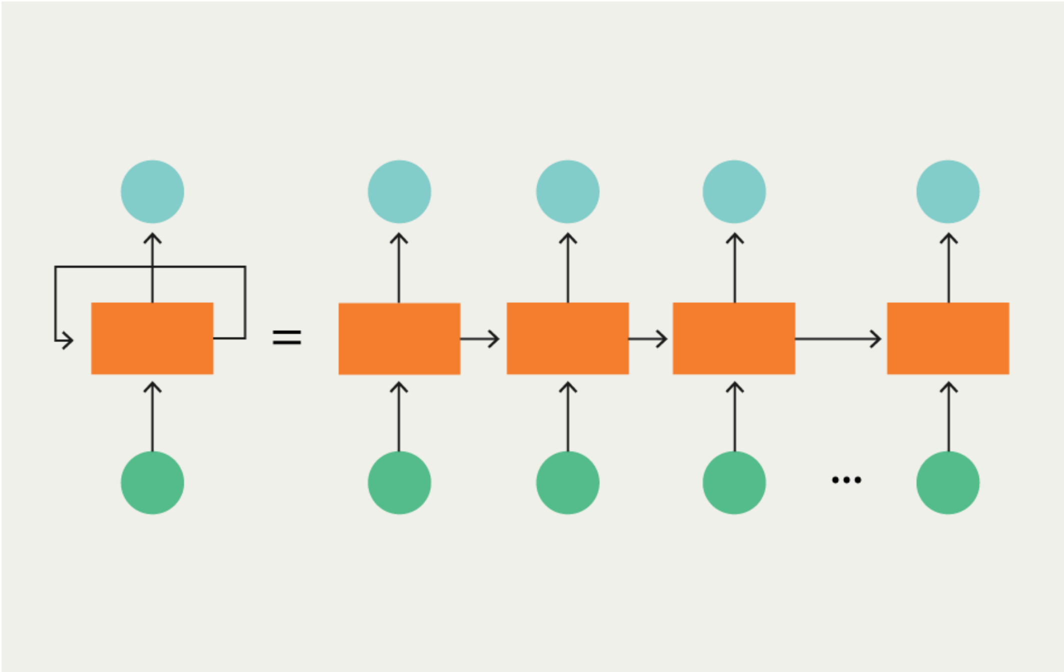

Figure 9: An RNN neural network [19]

Both in the Visual Cortex of the brain and in the neural networks, the output of each layer is the input of the next layer, and by having the output information of each layer, we can analyze and compare it with the layers of any model.

2-3 Comparison between Brain and Neural Networks with RSA

Representational Similarity Analysis or RSA for short is used to compare macaque's VC data with neural networks. RSA is a way to compare a group of numbers with another group of numbers that have different sizes. In this research, 352 of macaque's brain neurons, which were all part of the IT region of the Visual Cortex, have been examined. Each one has a value for each image, so for each photo, which we have a total of 81 photos, there are 352 at hand. By measuring the Euclidean distance of the 352 numbers of one photo and 352 numbers of another photo, a single number is obtained which can be interpreted as the distance of these two images from each other. In the Representational Dissimilarity Matrix, or abbreviated as RDM, rows and columns are the images that are used, which are a total of 81 images. Each cell is the distance between two specific images, and the diameter of the matrix is the distance of each picture from itself, which is equal to zero. In this way, each pair of images has a cell with a value of one number. Now, if we count these numbers for each double pair of these 81 images, we will obtain 3321 numbers. The distance of each photo from itself is zero, but the rest of the numbers are not necessarily zero. Also, the RDM matrix is a symmetrical matrix.

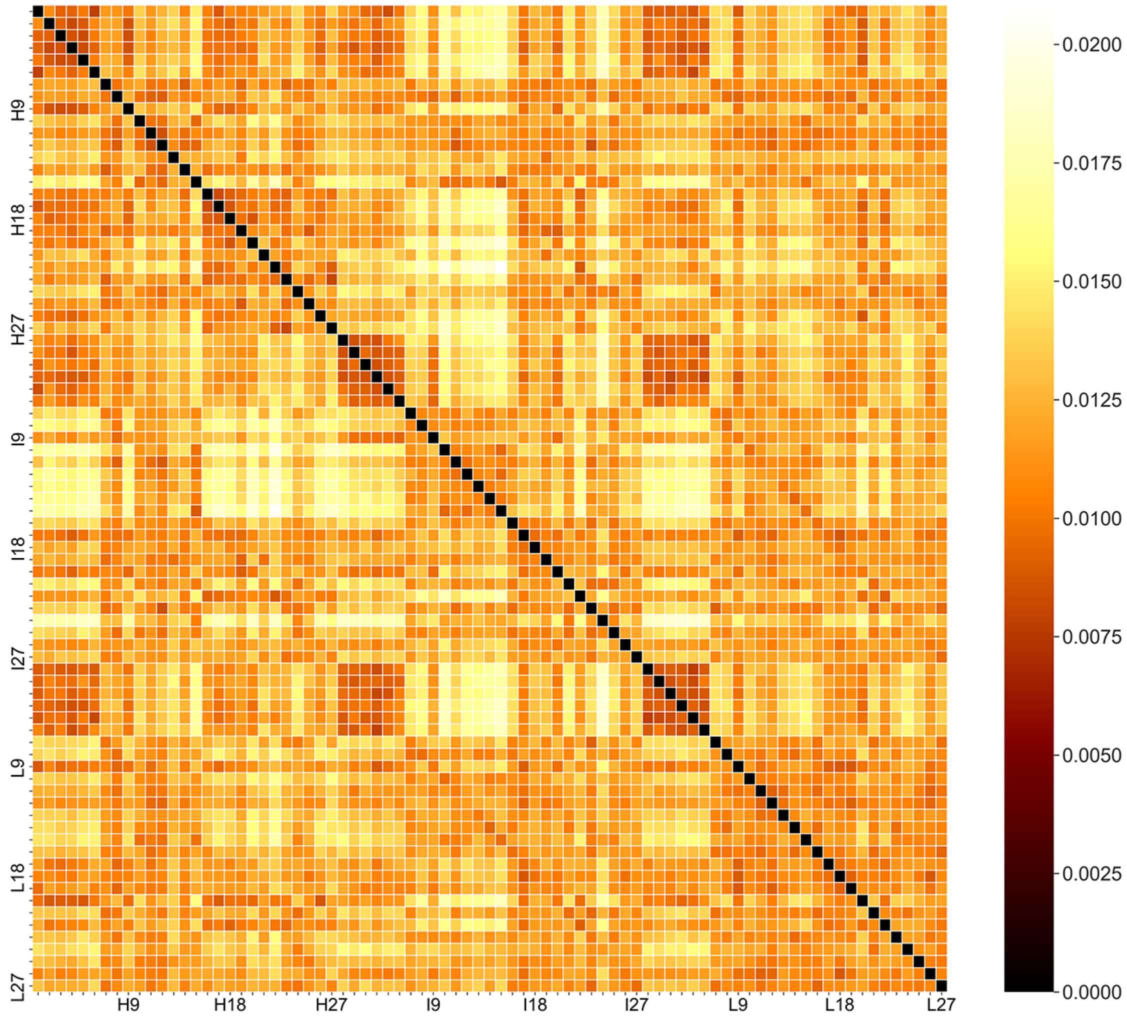

Figure 10: An example of an RDM, this RDM is for the IT region of macaque's VC. The rows and columns of this matrix are the images that have been used, which are a total of 81 photos.

The letter 'H' indicates sharp photos, the letter 'I' indicates Intact images, and the letter 'L' indicates blurry images.

Each cell is the distance of two images from each other. The diameter of the matrix is the distance of each image from itself, which is equal to zero. For a better understanding of the distances, the numbers in the cells have been normalized.

A similar process is carried out for each neural network. Now each two matrices can be compared. The way to compare the output of the VC and the layers of neural networks is by measuring the correlation between the RDMs of each layer of a neural network and the RDM of the IT region of the macaque's brain.

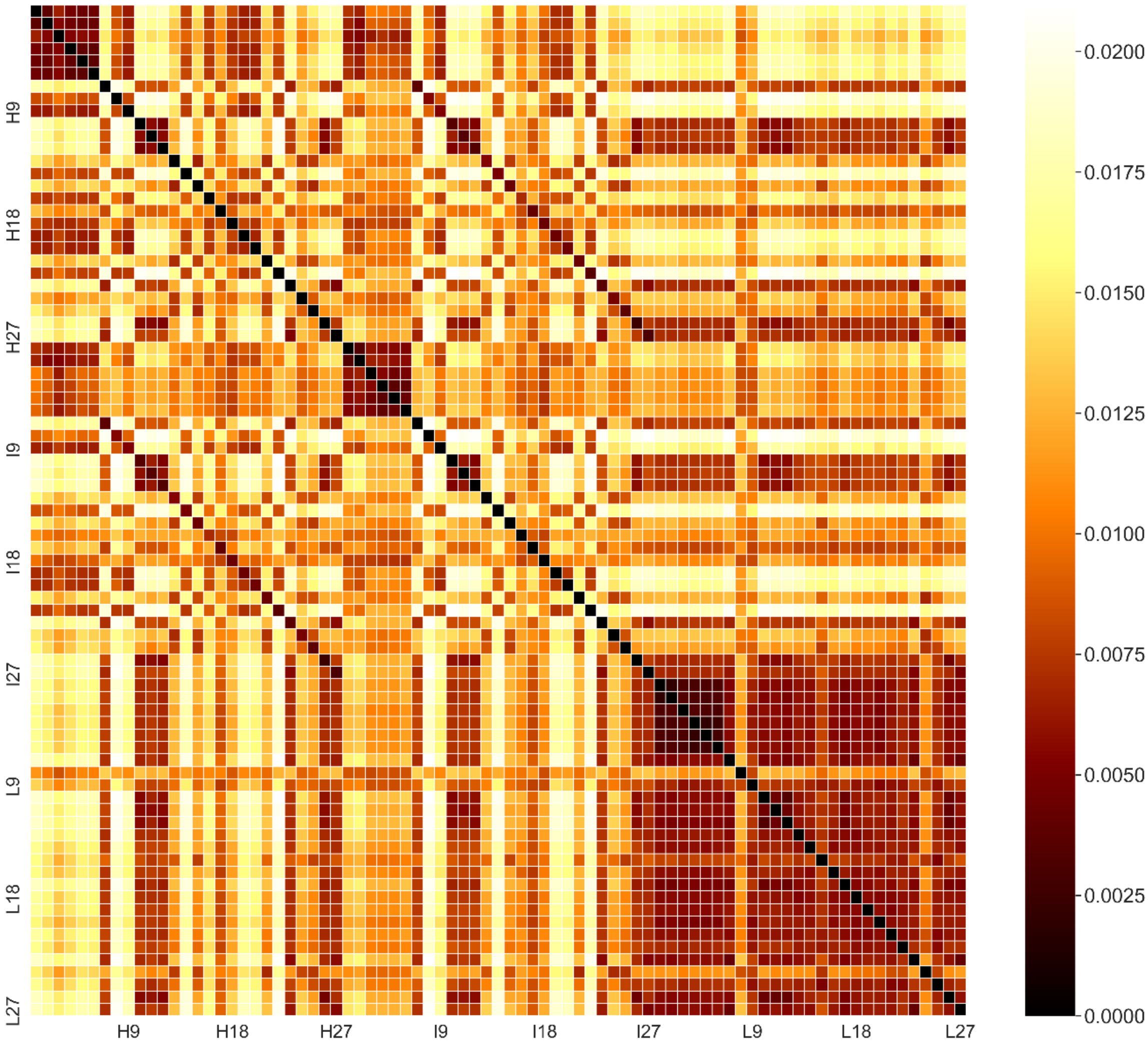

Figure 11: An example of an RDM, this RDM is for the ‘features.14.Conv0’ layer of the MobileNet-V2 neural network. The rows and columns of this matrix are the images used, which are a total of 81. The values of the cells are normalized.

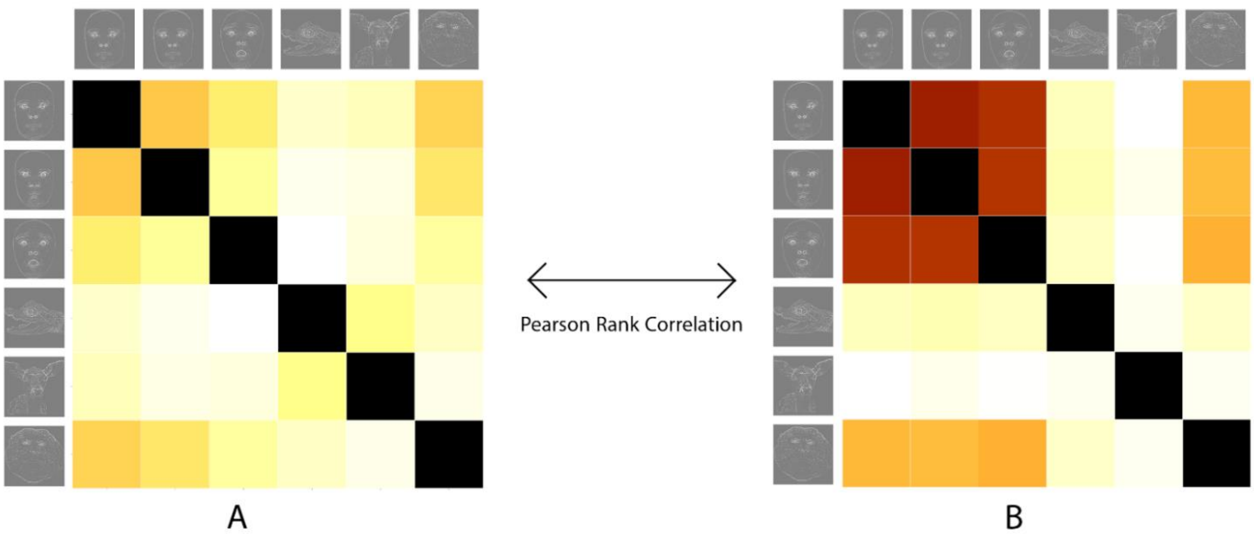

Figure 12: Comparison of two RDM matrices by calculating the Pearson Rank Correlation value of them.

Matrix A is an example of brain RDM matrix.

Matrix B is an example RDM matrix of one of the layers of ResNet18 neural network.

2-4 The Effect of Frequency on the Brain and Neural Networks

To compare the effects of frequency on the IT region of the brain with the neural networks, an SVM classifier is used. To train this SVM classifier, the data of the neurons for each of the intact, sharp and blurry images along with their labels is used, and to test this SVM classifier, the data of the neurons for other frequencies along with their labels is used. For neural networks, this process is done for the layer before the classifier of the network. The data of the layer before the classifier of the network, for each of the intact, sharp and blurry images along with their labels is used to train an SVM classifier, and to test it, the data of other frequencies along with their labels are used. By measuring the accuracy of this classifier, it is possible to compare the effect of frequency on neural networks and the IT region.

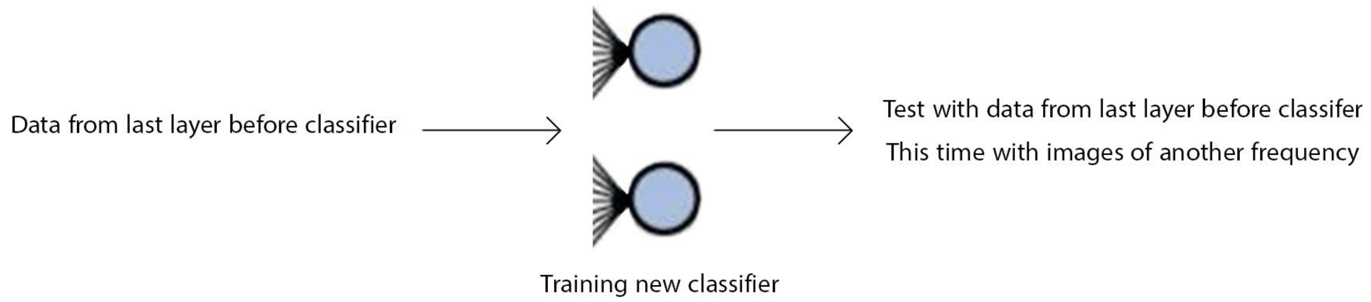

Figure 17: To analyze the effect of frequency on neural networks, a classifier is trained, whose input is the output of the layer before the classifier of the neural network for images of one frequency along with their labels. For example, the classifier was trained with the data of blurry images and the data of sharp images was tested on it.

2-5 Summary

In this chapter, we examined the structure of the Visual Cortex of macaque's brain and the function of its different regions. Then we explained how neural networks work and showed an example of the output of their layers for the images used in this research. Then we discussed how to compare the layers of neural networks and the IT region of the Visual Cortex of the brain through RSA analysis with the formation of RDM matrices. Finally, we discussed how to investigate the effect of frequency on the function of the Visual Cortex and neural networks.

Chapter 3:

Methodology

3-1 Outline

Now that we are familiar with the basic concepts needed, we can go ahead with the research. Our main goals are:

1. Comparing different layers of neural networks with the IT region of the Visual Cortex of macaque's brain and analyzing their similarities or differences.
2. Investigating the effect of the frequency of the images on the function of the IT region of the brain and the layer before the classifier layer of neural networks.

We follow these two objectives for AlexNet, GoogleNet, MobileNet-V2, ResNet18, ResNet50, Resnet101, VGG-16, VGG-19 and SqueezeNet1.1 neural networks.

3.2 Implementation

3-2-1-1 Brain Activity

In each session of measuring the activity of neurons in the IT region of macaque's brain, the activity level of each neuron was recorded for every millisecond, from 300 milliseconds before the image was shown to the monkey until 600 milliseconds after the image was shown. Then, to calculate the activity level of each neuron, the data starting from 70 milliseconds after displaying the image until 200 milliseconds after displaying the image has been averaged to obtain the activity of that neuron. Finally, the end result is the activity of 352 neurons of the IT region for 81 different images, of which 27 images are intact, 27 are sharp and 27 are blurry.

3-2-1-2 Comparing Neural Network Layers with the IT Region

This comparison is done by examining the degree of correlation between the RDMs of each layer of the neural network and the RDM of the IT region of the brain. First, the RDM matrix is calculated for the IT neurons. In this study, we have access to the data of 352 neurons from the IT region of the brain. Second, the RDM matrix is calculated for each layer of the neural network. Each layer of a neural network contains a large number of neurons, the number of which depends on the structure of each layer and the overall structure of the neural network. Third, between the RDM of each layer of each neural network and the RDM of the brain region, the Pearson Rank Correlation value is calculated and for each neural network, all the correlations are drawn as a graph.

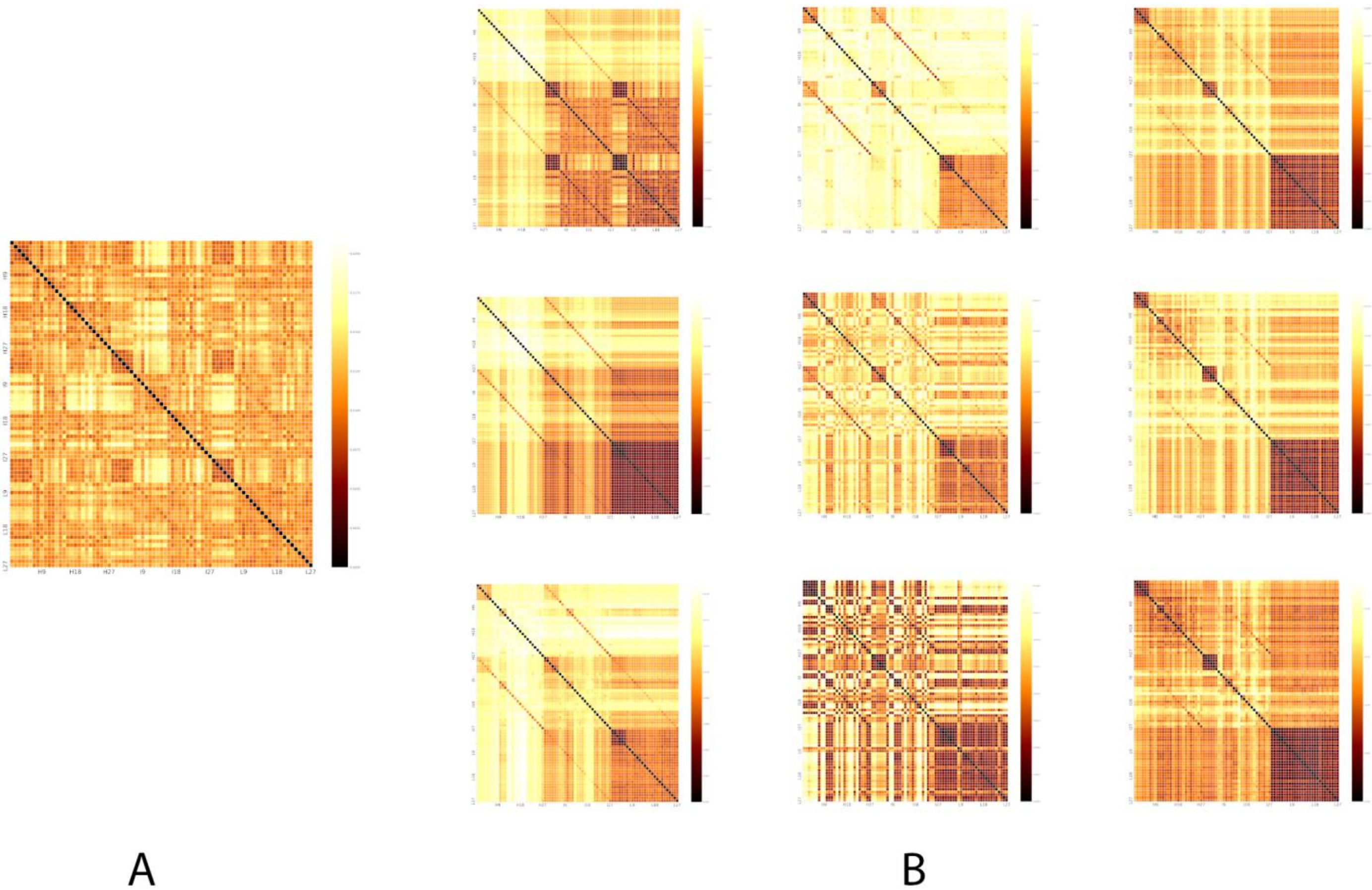

Figure 13: RDM matrices for the IT region of the brain and several different layers of the MobileNet-V2 neural network.

Figure A shows the RDM matrix for the IT region of the brain. The data has been normalized to make the comparison tangible.

Figure B shows the RDM of some layers of the MobileNet-V2 neural network, which is a CNN. Here, too, the data is normalized.

3-2-2 Effects of Frequency on Neural Networks and the Brain

For the brain, this is done by training a classifier with brain's data for images of one frequency with their labels and testing that classifier with brain's data for images of another frequency with their labels. For each neural network, this work is done by a classifier trained with the data of the layer before the classifier layer of the neural network for images of one frequency along with their labels and tested with the data of the layer before the classifier layer for images of another frequency along with their labels.

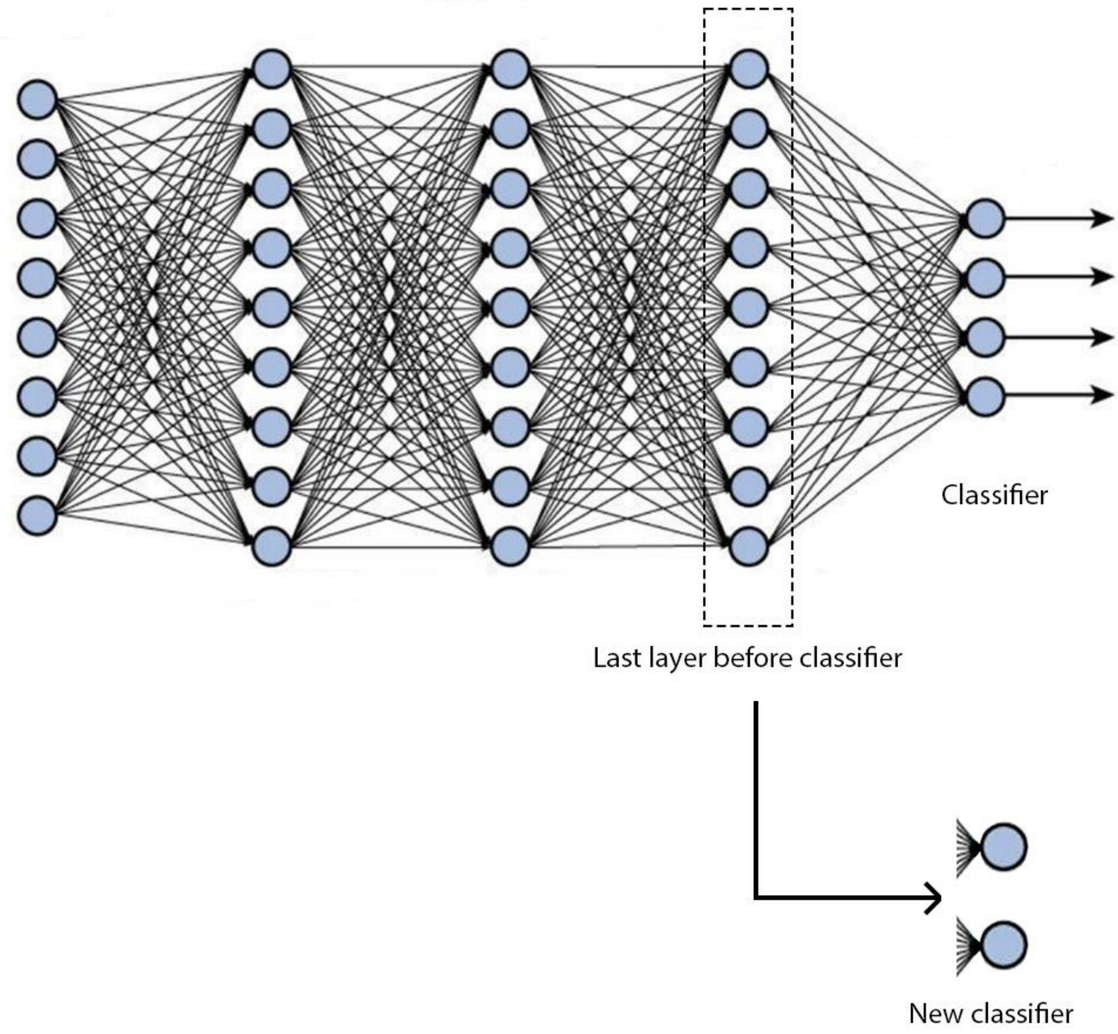

Figure 14: To analyze the effect of frequency on the neural network, a classifier is trained, whose input is the output of the layer before the classifier of the neural network for the images of a frequency along with their labels. For example, the classifier was trained with the data of blurry images and the data of sharp images was tested on it.

3-3 Summary

In this chapter, we explained how to carry out the study and the methods for comparing the brain with neural networks.

Chapter 4:

Implementation and Results

4.1 Results

4-1-1 Comparison of the IT Region with layers of neural networks

In this section, we compare the IT region of the brain with all layers of AlexNet, GoogleNet, MobileNet-V2, ResNet18, ResNet50, Resnet101, VGG-16, VGG-19 and SqueezeNet1.1 neural networks. As it was stated, this is done by calculating the correlation value of RDMs of each layer of the neural networks and the IT region. Here we deal with face and body images. Face images include 3 human face images and 3 animal face images, each in 3 different frequencies. Also, the body images include 3 human body images and 3 animal body images, each in 3 different frequencies.

In the following, for each neural network, a face correlation diagram in 3 different frequencies, a body correlation diagram in 3 different frequencies, a best layer diagram for face and body in 3 different frequencies and a maximum correlation diagram in 3 different frequencies are drawn. For the correlation diagrams, the horizontal axis from left to right are the initial and final layers of the neural network, respectively.

If the final layers of the neural network are comparable to the IT region, we expect the correlation value to increase as we move from the initial layers of the neural network to the final layers of the neural network. We do this comparison in 3 different frequencies. In the graphs, sharp images are shown in red, intact images are shown in blue, and blurry images are shown in green.

Important Note:

In this research, low frequency images and blurry images are interchangeable phrases. Intact frequency images and intact images are interchangeable phrases. High frequency images and sharp images are also interchangeable phrases.

4-1-1-1 AlexNet

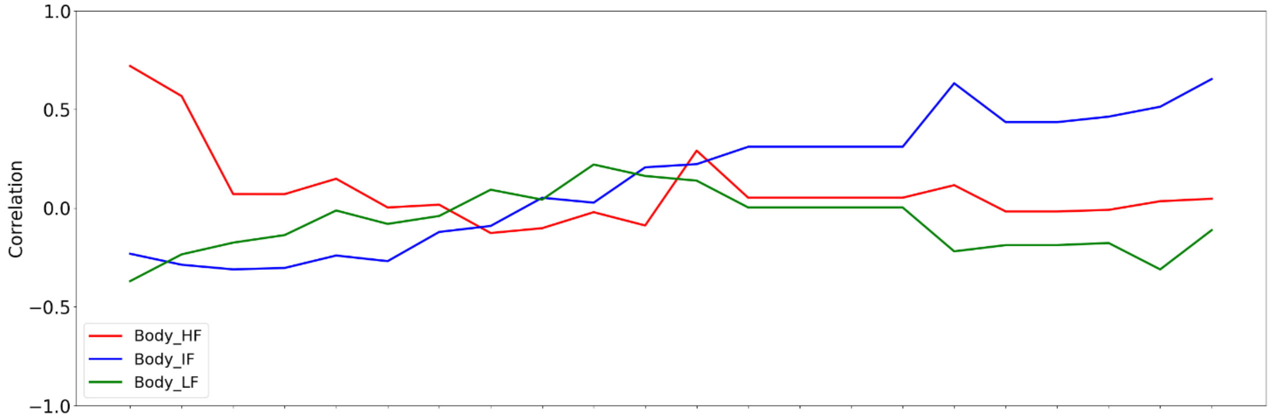

Figure 15: Correlation diagram of RDMs of AlexNet network and the brain for body images in intact, high and low frequencies.

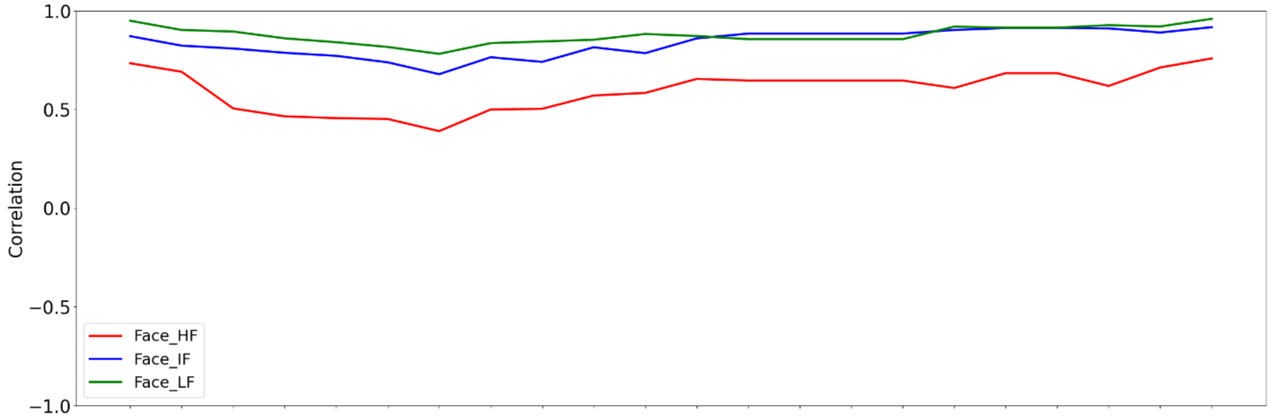

Figure 16: Correlation diagram of RDMs of AlexNet network and the brain for face images in intact, high and low frequencies.

For body images, the trend of each frequency is different from the other. In the initial layers, the blurry and intact images behave the same, but in the higher layers, the blurry and sharp images are comparable. Also, only for the intact body images, the similarity with the IT region of the brain increases as the later layers are approached, but for other frequencies. In the face images, the correlations of all the three frequencies are comparable to each other and their trends are static, but the blurry and intact images are more comparable. As a result, for face images, the last layers are not comparable to the IT region of the brain more than the initial layers.

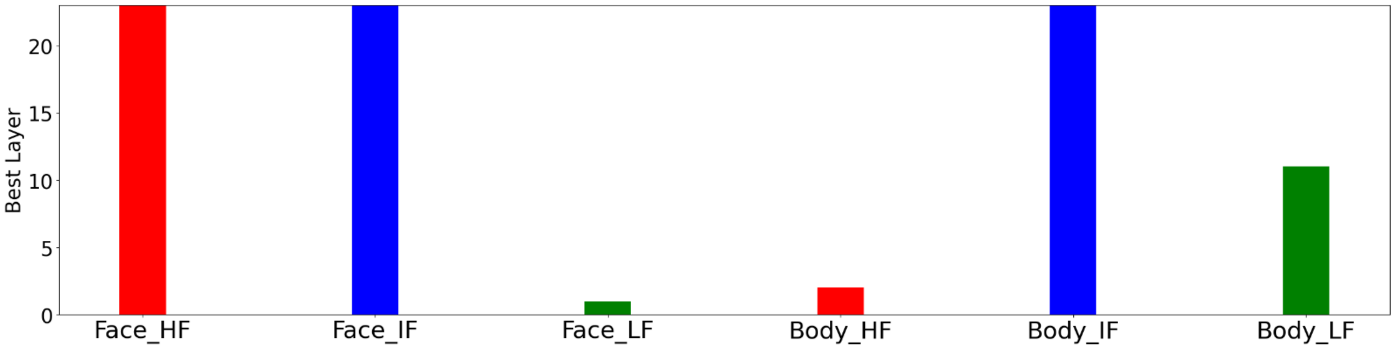

Figure 17: AlexNet layers that have the highest correlation with the brain for face and body images in intact, high and low frequencies.

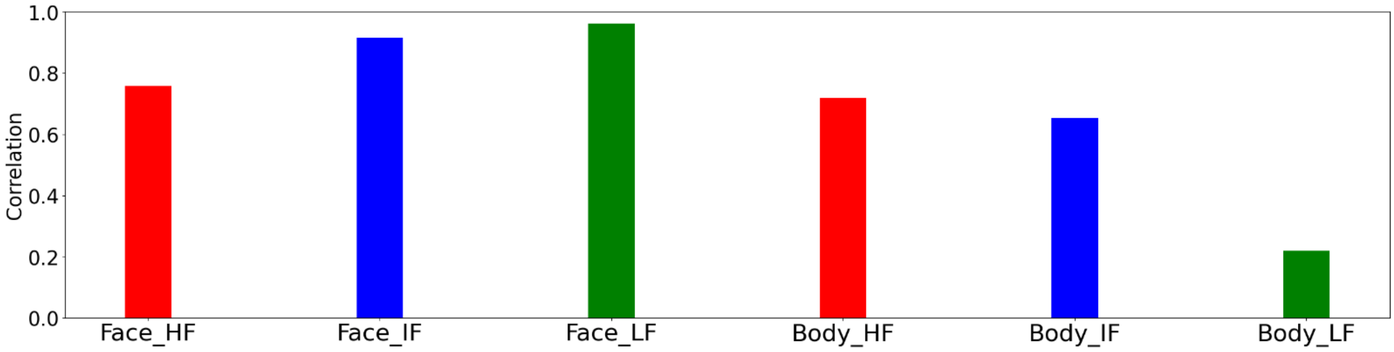

Figure 18: The highest correlation between the layers of the AlexNet network and the brain for face and body images in in intact, high and low frequencies.

According to the above 2 diagrams, for some images such as intact face and body images and sharp face images, the last layers are more comparable, and the first layers are better for blurry face images and sharp body images.

Also, the maximum correlation for each category of images has been shown.

4-1-1-2 GoogleNet

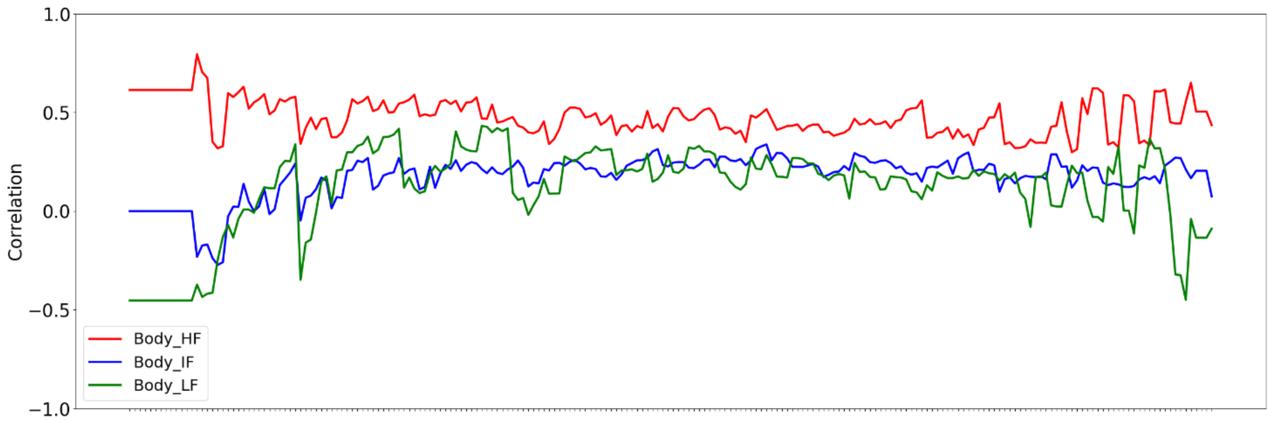

Figure 19: Correlation diagram of GoogleNet and brain RDMs for body images in intact, high and low frequencies.

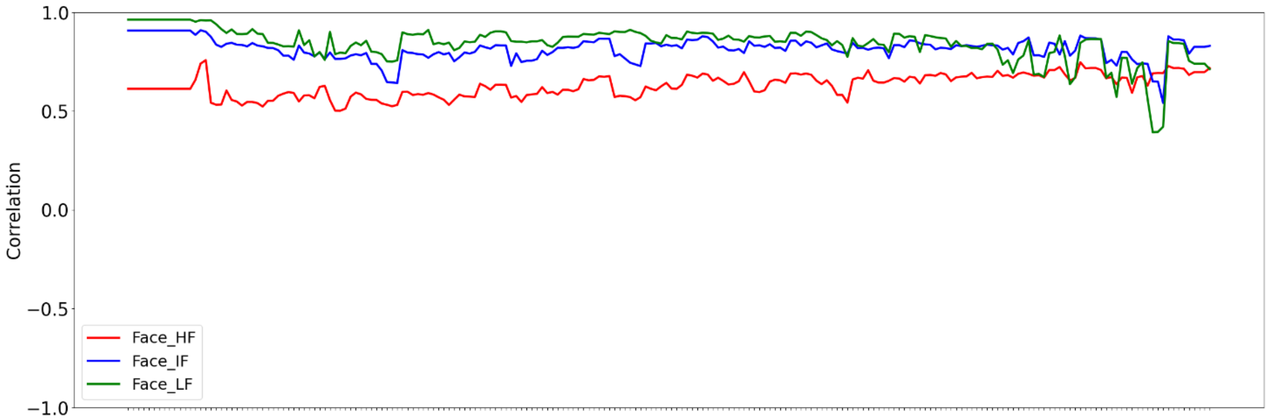

Figure 20: Correlation diagram of GoogleNet and brain RDMs for face images in intact, high and low frequencies.

For body images, the general trends of intact and blurry images are the same, but sharp images behave differently. Also, as the final layers are approached, the similarity with the IT region of the brain does not increase, in any of the frequencies.

For face images, the correlation of all three frequencies does not increase, and their trends are similar, but the blurry and intact images are more comparable. blurry images also have a negative trend at the end. As a result, for face images, the last layers are not comparable to the IT region of the brain more than the first layers.

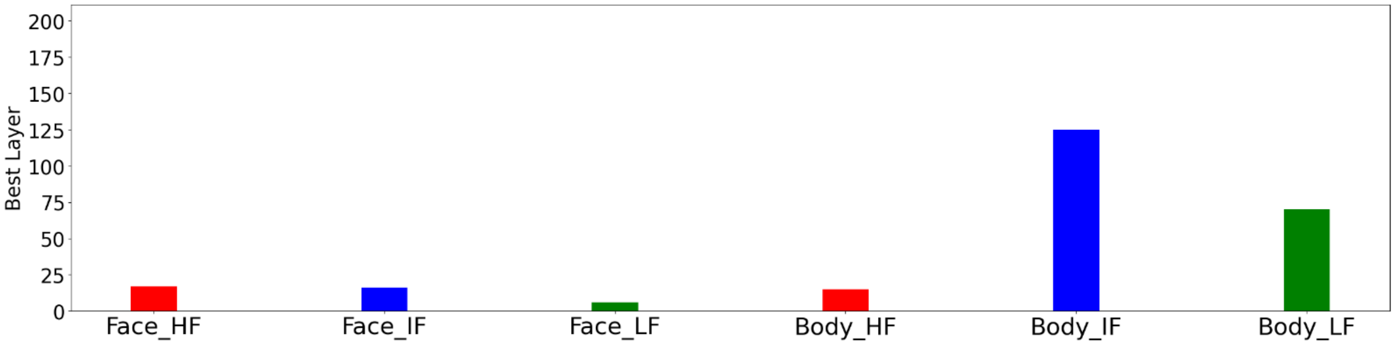

Figure 21: The layers of the GoogleNet network that have the highest correlation with the brain for face and body images in intact, high and low frequencies.

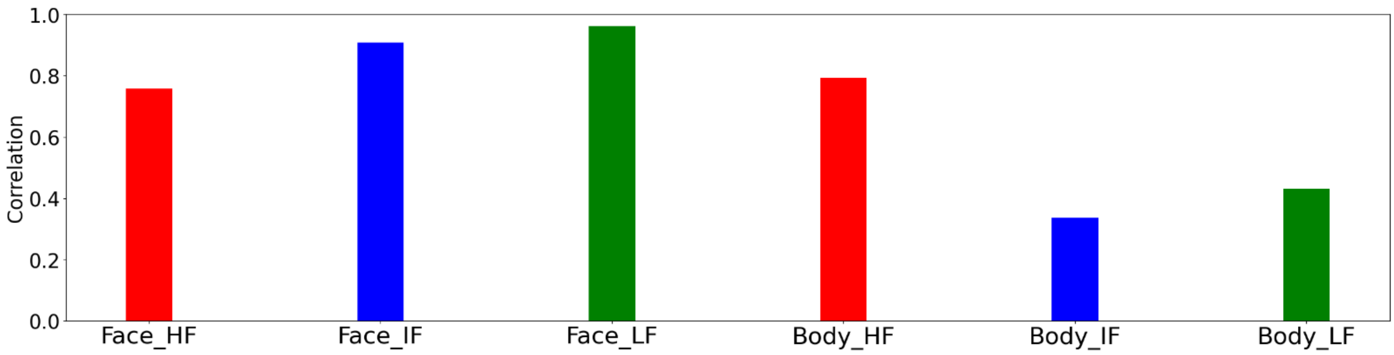

Figure 22: The highest correlation between GoogleNet and the brain for face and body images in intact, high and low frequencies.

According to the 2 graphs above, the best layers for most images are the initial layers, and only for intact body images, the middle layers are more comparable.

Also, the maximum correlation for each category of images has been shown.

4-1-1-3 MobileNet-V2

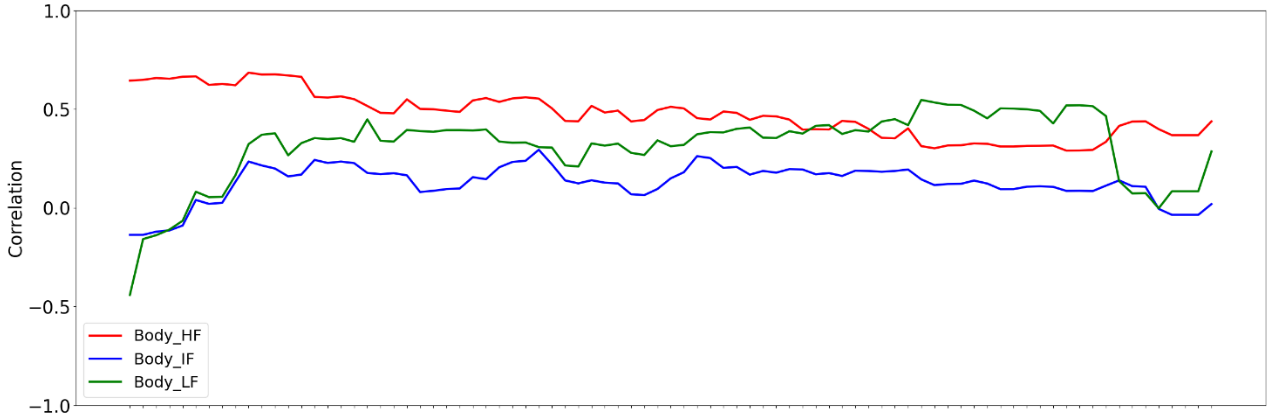

Figure 23: Correlation diagram of RDMs of MobileNet-V2 network and the brain for body images in intact, high and low frequencies.

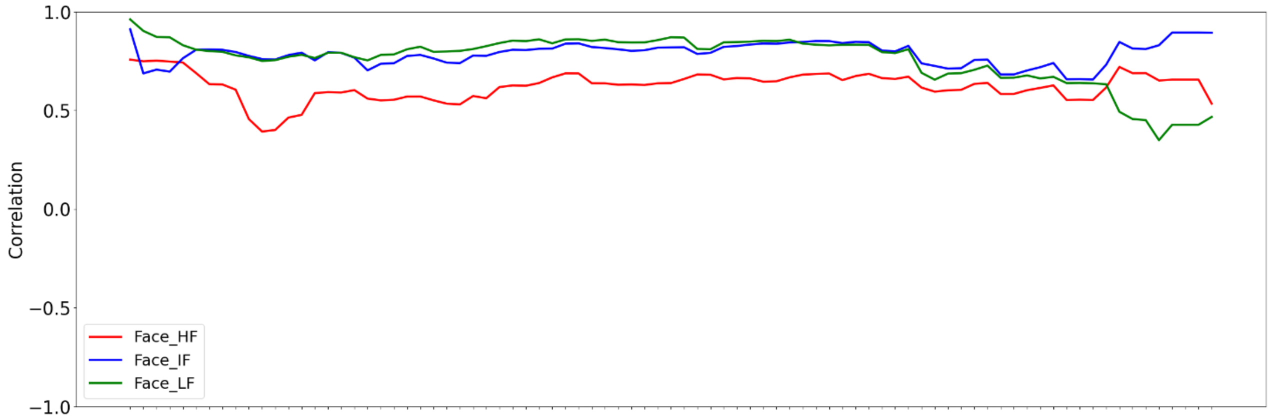

Figure 24: Correlation diagram of RDMs of MobileNet-V2 network and the brain for face images in intact, high and low frequencies.

For body images, the general trends of intact and blurry images are the same, although the trend of blurry images changes a little at the end, sharp images behave differently. Also, as the final layers are approached, the comparability with the IT region of the brain does not increase in any of the frequencies.

In face images, the correlations for all three frequencies are similar to each other and their trends are static, but the blurry and intact images are more comparable. blurry images also have a negative trend at the end. As a result, for face images, the last layers are not comparable to the IT region of the brain more than the initial layers.

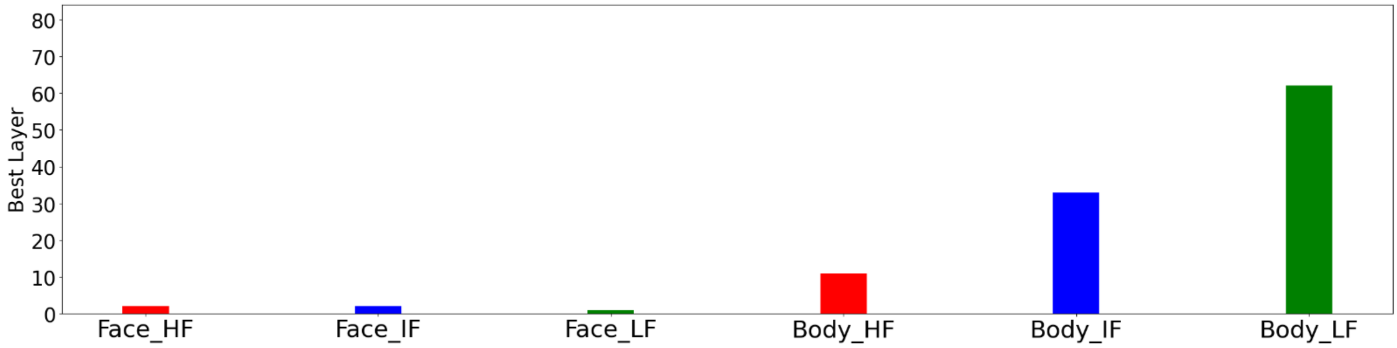

Figure 25: The layers of the MobileNet-V2 network that have the highest correlation with the brain for face and body images in intact, high and low frequencies.

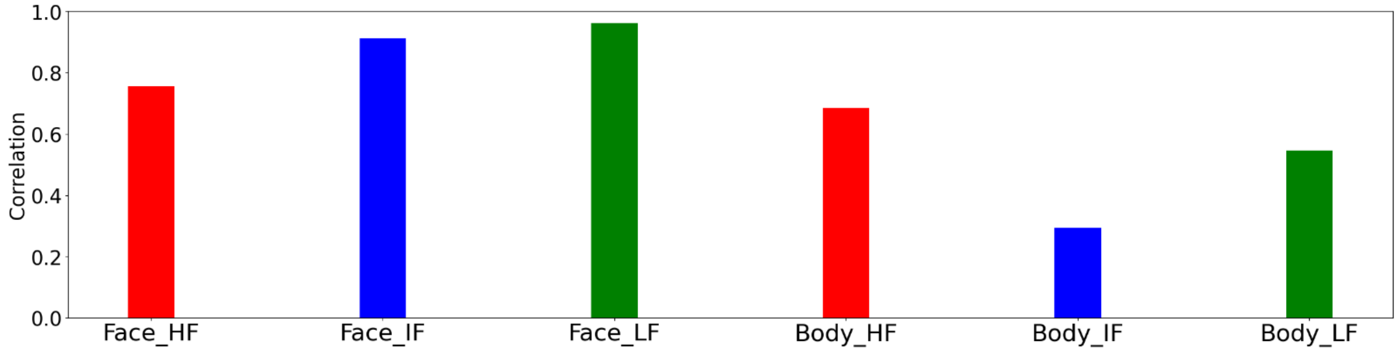

Figure 26: The highest correlation between MobileNet-V2 layers and the brain for face and body images in intact, high and low frequencies.

According to the above 2 diagrams, the best layers for most images are the initial layers, but for intact body images, the middle layers are more comparable, and for blurry body images, the last layers.

Also, the maximum correlation for each category of images has been shown.

4-1-1-4 ResNet18

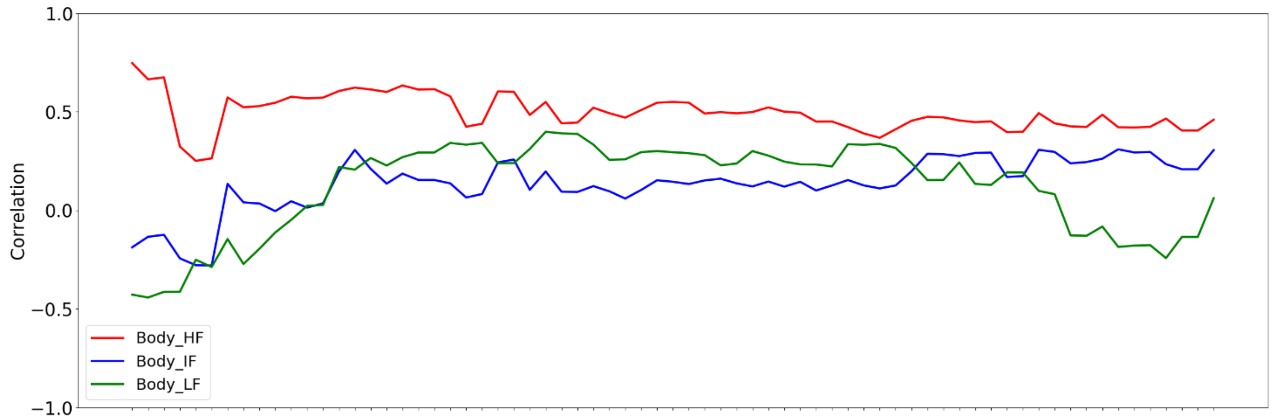

Figure 27: Correlation diagram of RDMs of ResNet18 network and the brain for body images in intact, high and low frequencies.

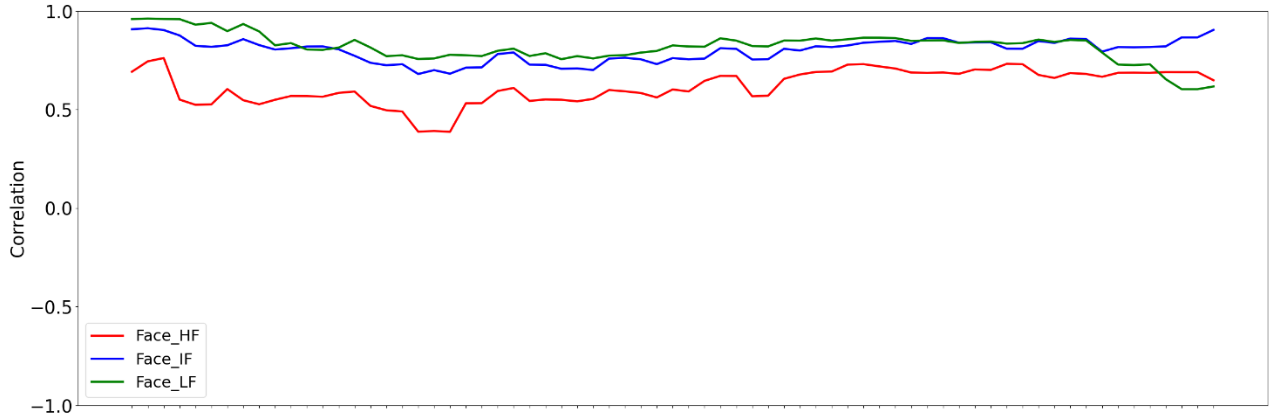

Figure 28: Correlation diagram of RDMs of ResNet18 network and brain for face images in intact, high and low frequencies.

For body images, the general trends of intact and blurry images are the same, although the trend of blurry images changes a little at the end. However, Sharp images behave differently. Also, as the final layers are approached, the comparability with the IT region of the brain does not increase in any of the frequencies.

In face images, the correlations for all three frequencies are comparable to each other and their trends are static, but the blurry and intact images are more comparable. blurry images also have a downward trend at the end. As a result, for face images, the last layers are not more similar to the IT region of the brain than the initial layers.

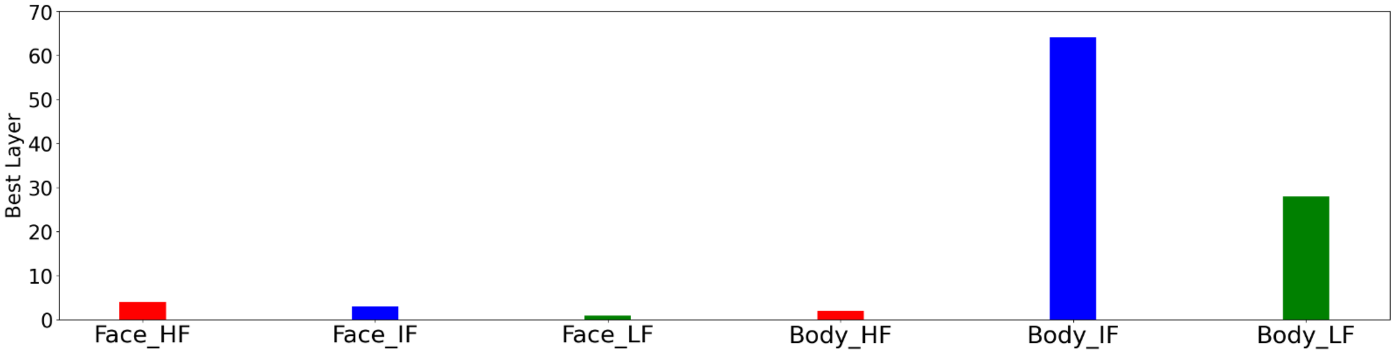

Figure 34: Layers of the ResNet18 network that have the highest correlation with the brain for face and body images in intact, high and low frequencies.

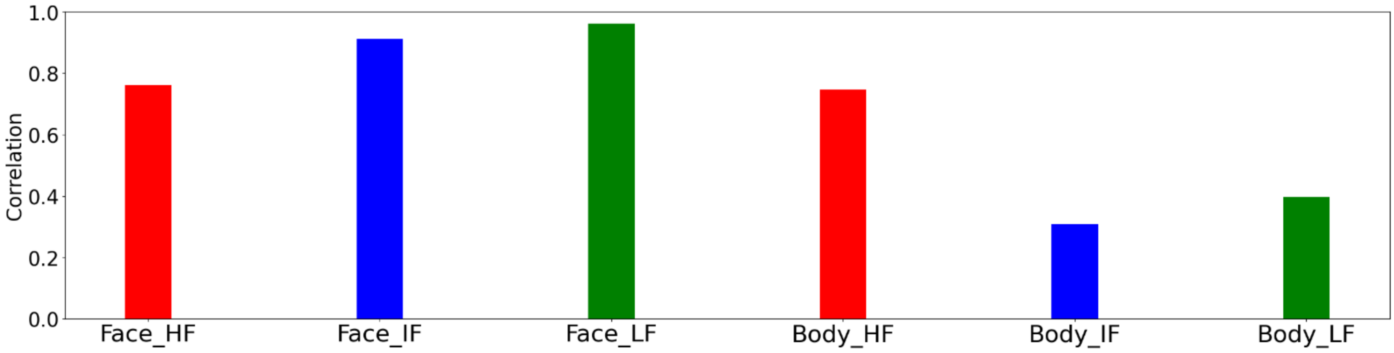

Figure 30: The highest correlation between ResNet18 layers and the brain for face and body images in intact, high and low frequencies.

According to the above 2 diagrams, the best layers for most images are the initial layers, but for intact body images, the final layers are more comparable, and for blurry body images, the middle layers.

Also, the maximum correlation for each category of images has been shown.

4-1-1-5 ResNet50

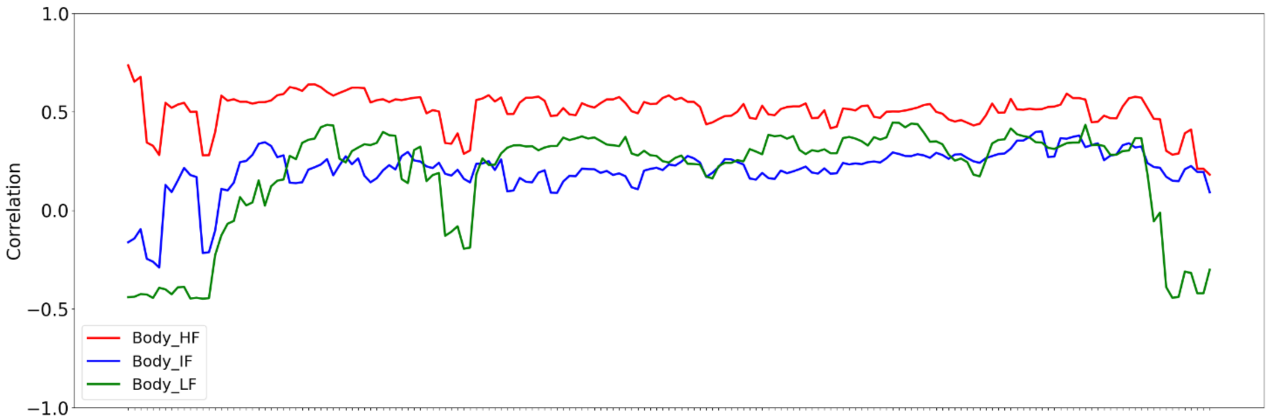

Figure 31: Correlation diagram of RDMs of ResNet50 network and the brain for body images in intact, high and low frequencies.

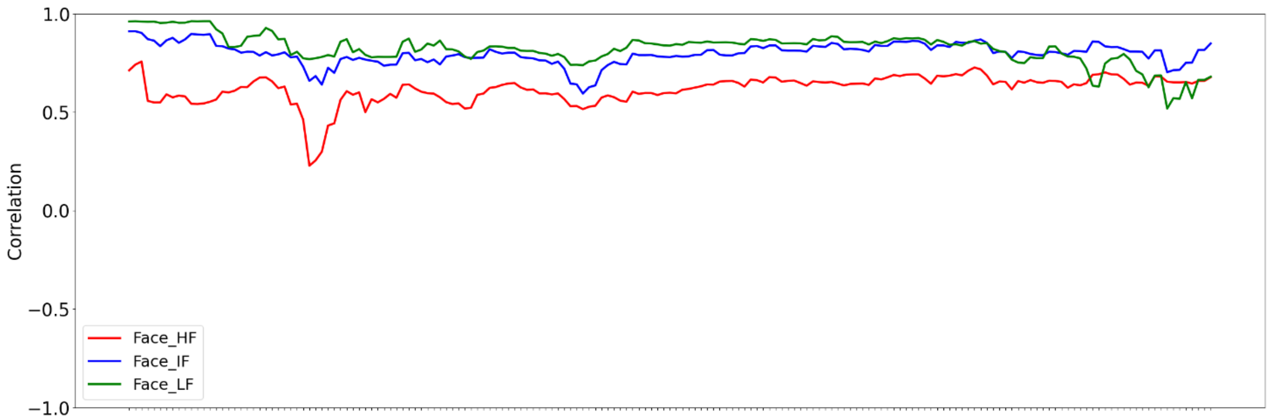

Figure 32: Correlation diagram of RDMs of ResNet50 network and the brain for face images in intact, high and low frequencies.

For body images, the general trends of intact and blurry images are the same, although the trend of blurry images changes a little at the end. However, Sharp images behave differently. Also, as the final layers are approached, the comparability with the IT region of the brain does not increase in any of the frequencies.

In face images, the correlations for all three frequencies are similar to each other and their trends are static, but the blurry and intact images are more comparable. blurry images also have a negative trend at the end. As a result, for face images, the last layers are not more similar to the IT region of the brain than the initial layers.

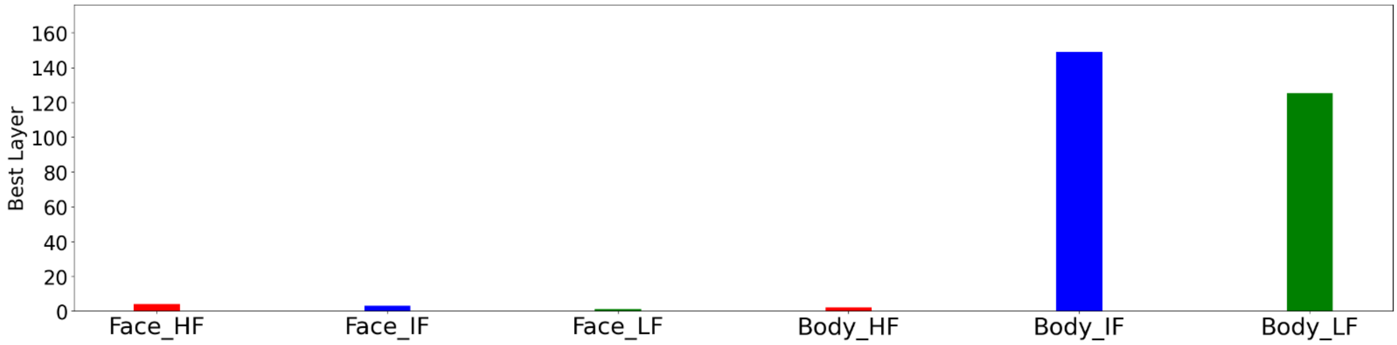

Figure 33: Layers of the ResNet50 network that have the highest correlation with the brain for face and body images in intact, high and low frequencies.

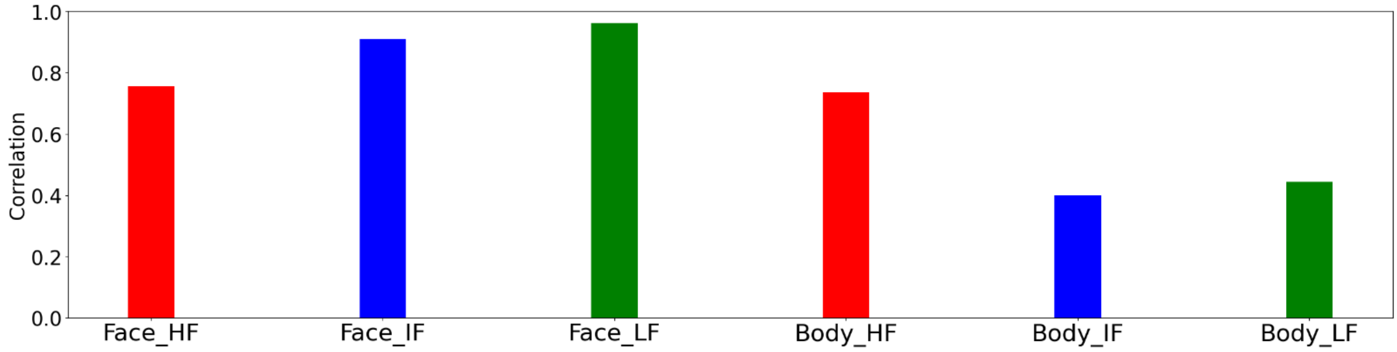

Figure 34: The highest correlation between the layers of the ResNet50 network and the brain for face and body images in intact, high and low frequencies.

According to the above 2 diagrams, the best layers for most images are the initial layers, but for intact and blurry body images, the final layers are more comparable.

Also, the maximum correlation for each category of images has been shown.

4-1-1-6 ResNet101

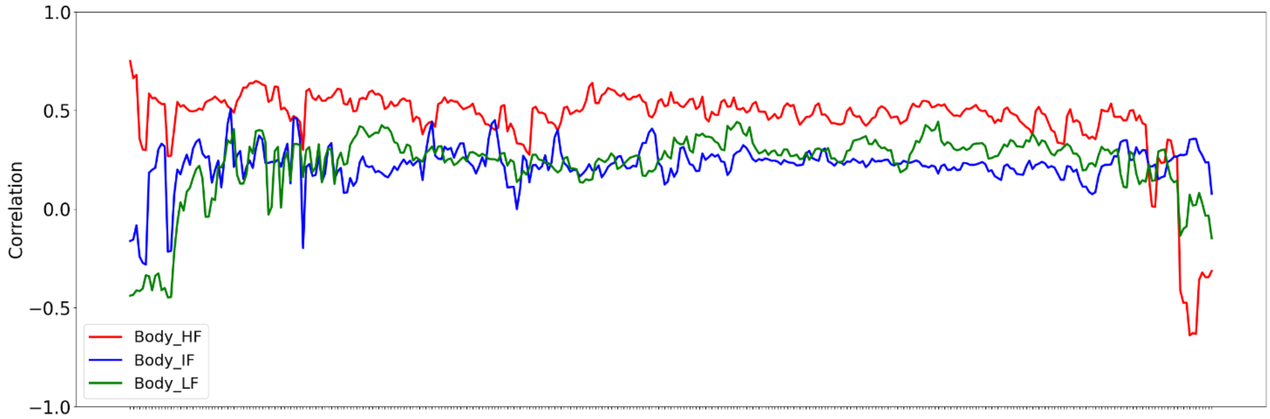

Figure 35: Correlation diagram of RDMs of ResNet101 network and the brain for body images in intact, high and low frequencies.

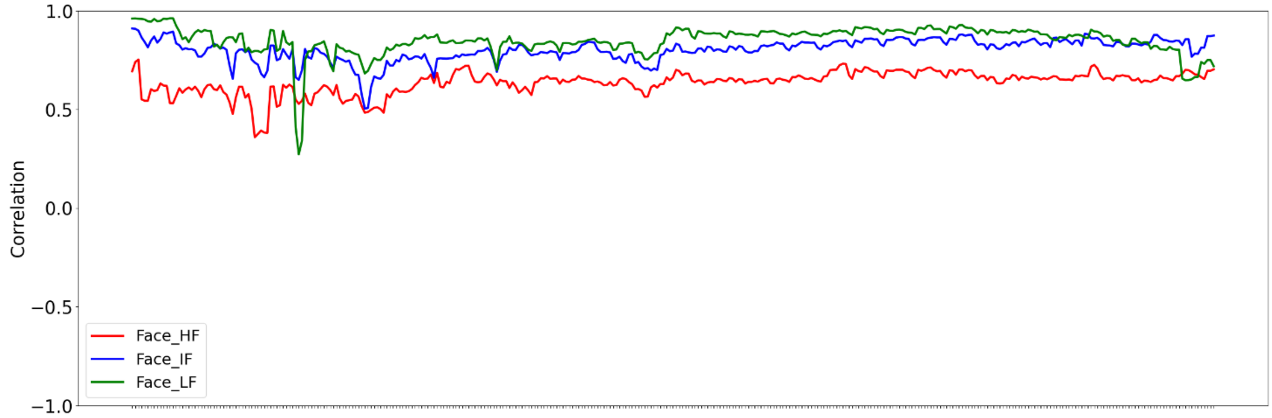

Figure 36: Correlation diagram of RDMs of ResNet101 network and the brain for face images in intact, high and low frequencies.

For body images, the general trends of intact and blurry images are the same, although the trend of blurry images changes a little at the end. However, sharp images behave differently. Also, as the final layers are approached, the comparability with the IT region of the brain does not increase in any of the frequencies.

In face images, the correlations for all three frequencies are similar to each other and their trends are static, but the blurry and intact images are more similar. blurry images also have a downward trend at the end. As a result, for face images, the last layers are not more comparable to the IT region of the brain than the first layers.

Figure 37: Layers of the ResNet101 network that have the highest correlation with the brain for face and body images in intact, high and low frequencies.

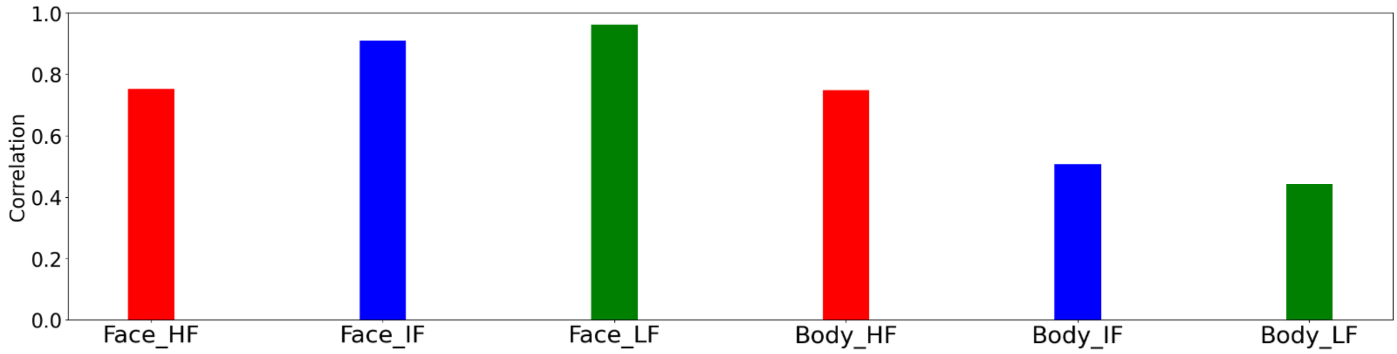

Figure 38: The highest correlation between the layers of the ResNet101 network and the brain for face and body in intact, high and low frequencies.

According to the above 2 charts, the best layers for most images are the initial layers, but for blurry body images, the end layers are more similar.

Also, the maximum correlation for each category of images has been shown.

4-1-1-7 SqueezeNet1.1

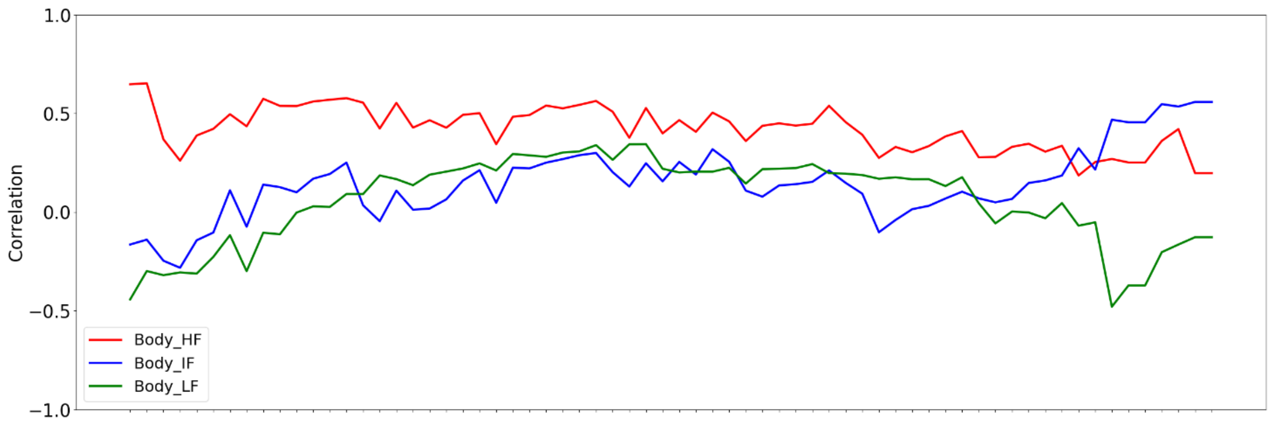

Figure 39: Correlation diagram of SqueezeNet1.1 and the brain RDMs for body images in intact, high and low frequencies.

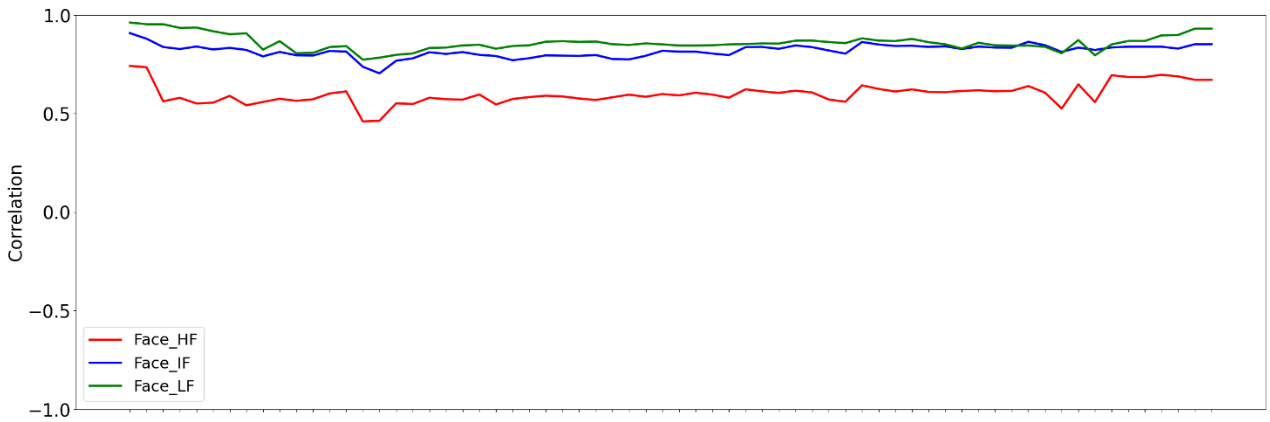

Figure 40: Correlation diagram of RDMs of SqueezeNet1.1 and the brain for face images in intact, high and low frequencies.

For body images, the general trends of intact and blurry images are the same, although the trend of matte images changes a little at the end, but sharp images behave differently. Also, only in intact images, the similarity with the IT region of the brain increases as the final layers are approached.

In face images, the correlations for all three frequencies are similar to each other and their trends are static, but the blurry and intact images are more similar. blurry images also have a downward trend at the end. As a result, for face images, the last layers are not more comparable to the IT region of the brain than the first layers.

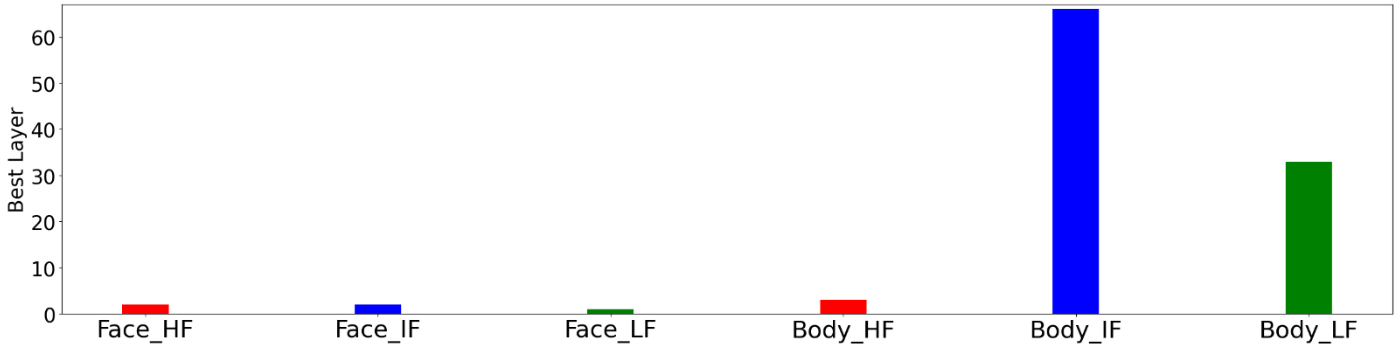

Figure 41: The layers of the SqueezeNet1.1 network that have the highest correlation with the brain for face and body images in intact, high and low frequencies.

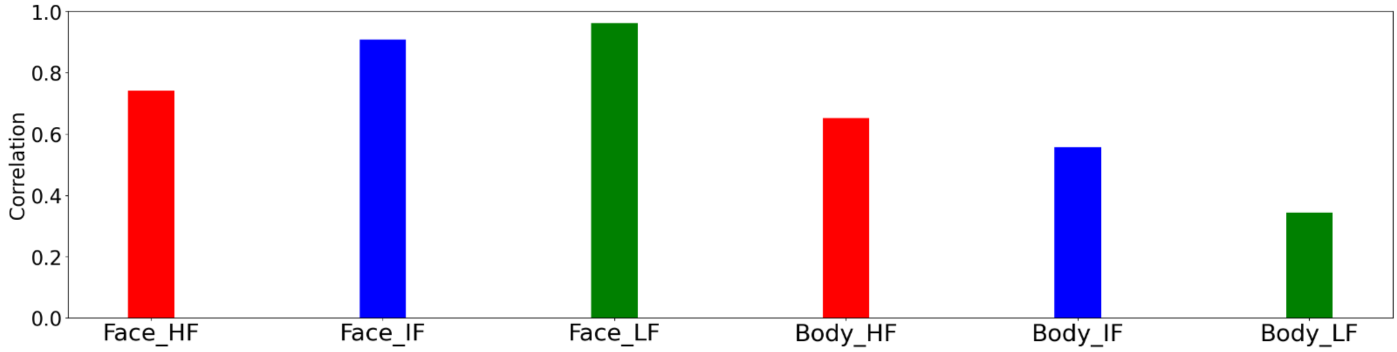

Figure 42: The highest correlation between SqueezeNet1.1 network layers and the brain for face and body images in intact, high and low frequencies.

According to the above 2 diagrams, the best layers for most images are the initial layers, but for intact body images, the final layers are more comparable, and for blurry body images, the middle layers.

Also, the maximum correlation for each category of images has been shown.

4-1-1-8 VGG-16

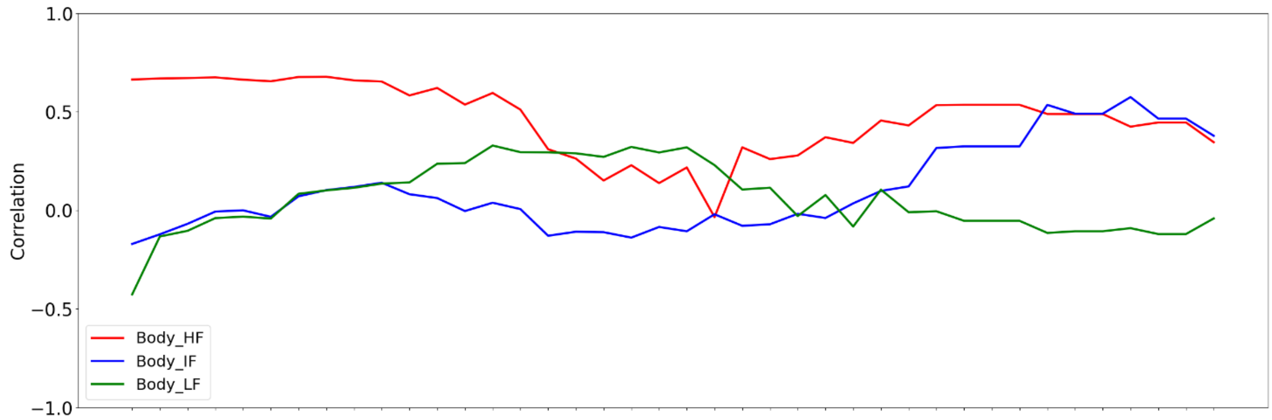

Figure 43: Correlation diagram of RDMs of VGG-16 network and the brain for body images in intact, high and low frequencies.

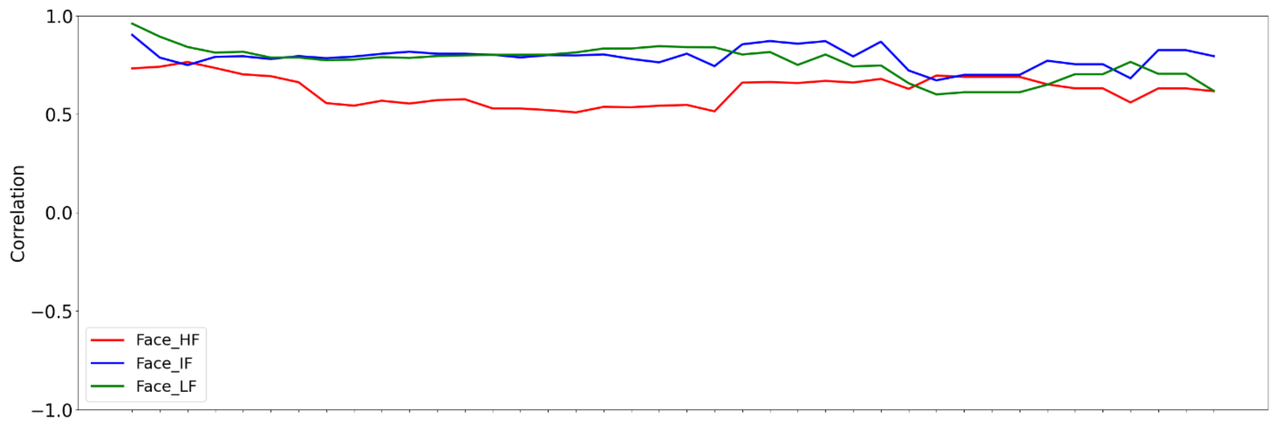

Figure 44: Correlation diagram of RDMs of VGG-16 network and brain for face images in intact, high and low frequencies.

For body images, each type of images behaves separately and none of them have similarities in the process. Also, only in intact images, the similarity with the IT region of the brain increases as the terminal layers are approached.

In face images, the correlations for all three frequencies are similar to each other and their trends are static, but the blurry and intact images are more similar. blurry images also have a downward trend at the end. As a result, for face images, the last layers are not more comparable to the IT region of the brain than the first layers.

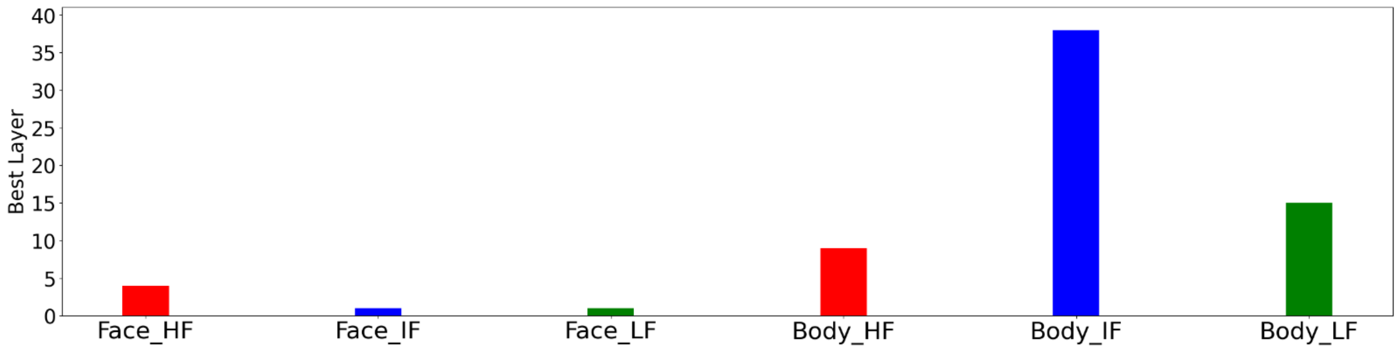

Figure 45: VGG-16 network layers that have the highest correlation with the brain for face and body images in intact, high and low frequencies.

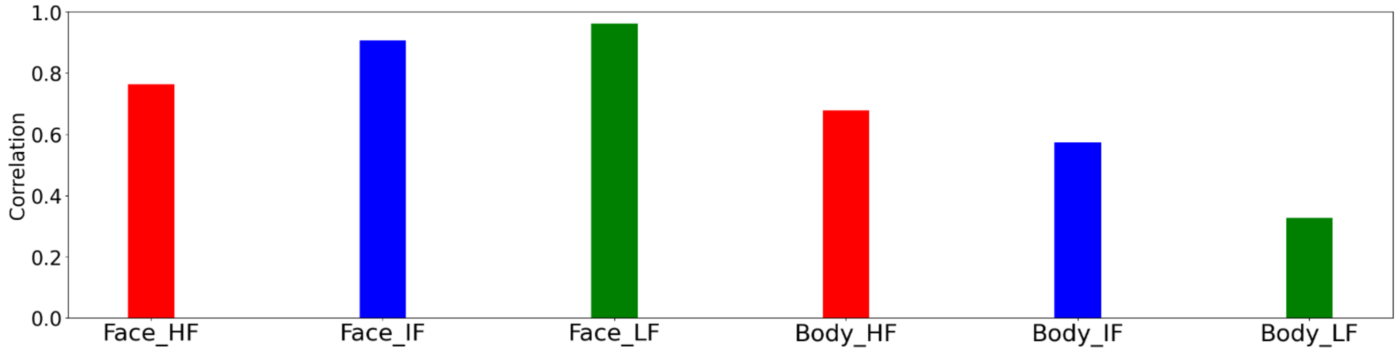

Figure 46: The highest correlation between VGG-16 network layers and the brain for face and body images in intact, high and low frequencies.

According to the above 2 diagrams, the best layers for most images are the initial layers, but for intact body images, the end layers are more similar, and for blurry body images, the middle layers.

Also, the maximum correlation for each category of images has been shown.

4-1-1-9 VGG-19

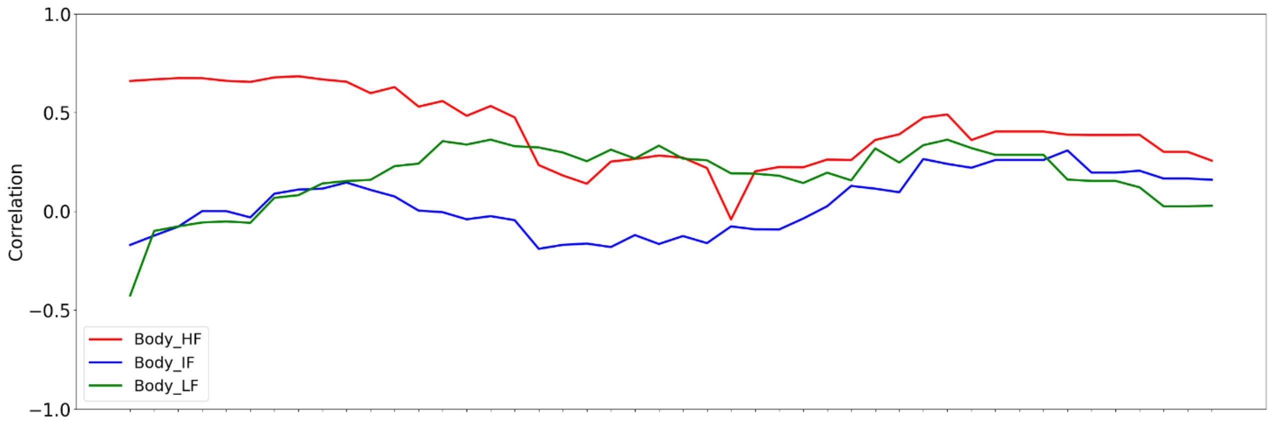

Figure 47: Correlation diagram of RDMs of VGG-19 network and the brain for body images in intact, high and low frequencies.

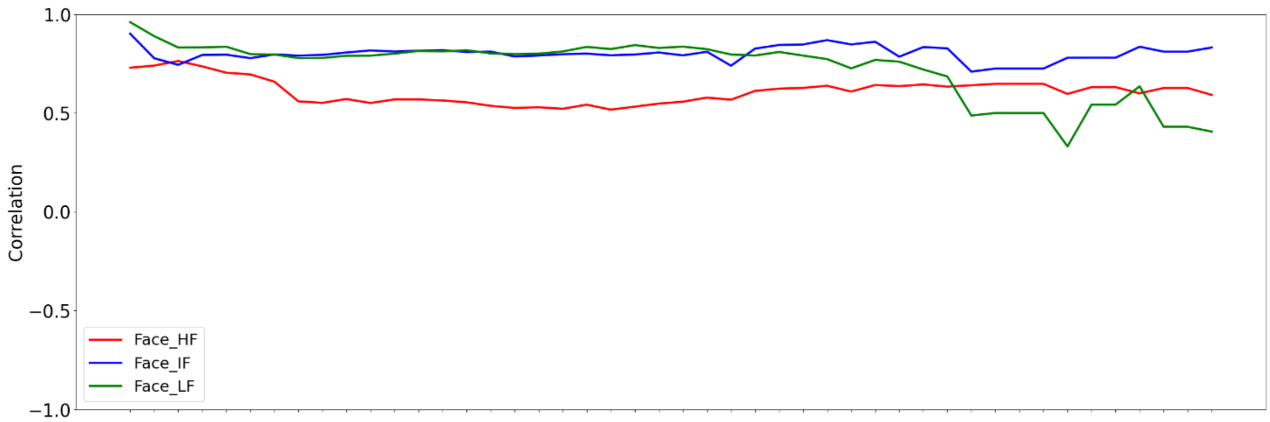

Figure 48: Correlation diagram of RDMs of VGG-19 network and brain for face images in in intact, high and low frequencies.

For body images, each type of images behaves differently and none of them have similarities in the process. Also, only in intact images, the similarity with the IT region of the brain increases as the final layers are approached.

In face images, the correlations for all three frequencies are similar to each other and their trends are static, but the blurry and intact images are more similar. blurry images also have a downward trend at the end. As a result, for face images, the last layers are not more comparable to the IT region of the brain than the first layers.

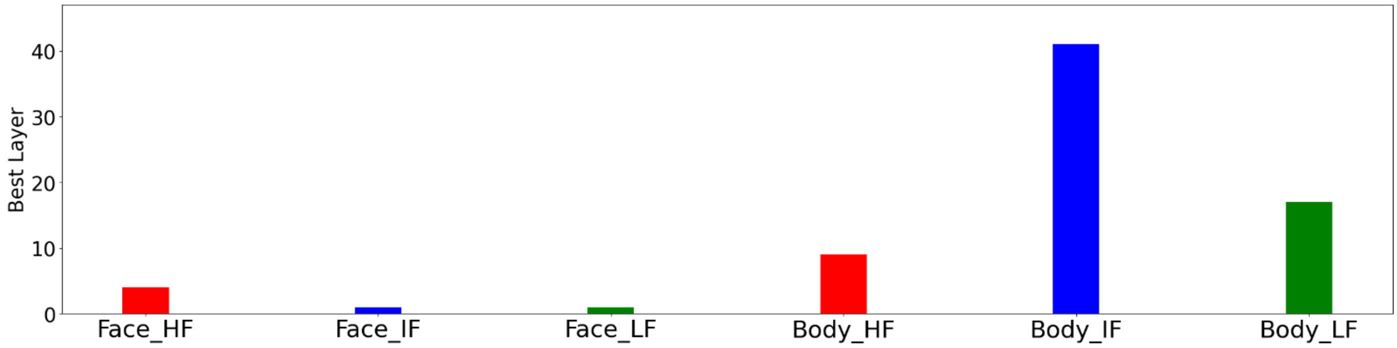

Figure 49: The layers of the VGG-19 network that have the highest correlation with the brain for face and body images in intact, high and low frequencies.

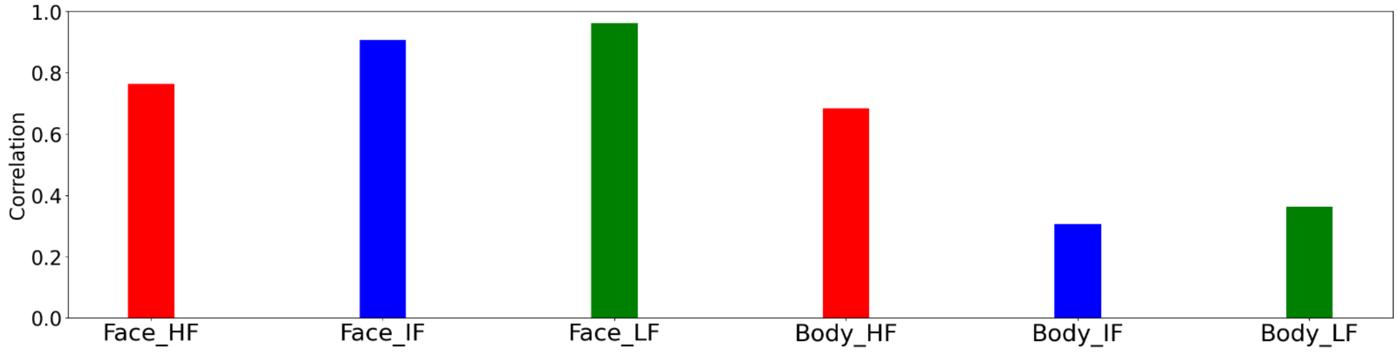

Figure 50: The highest correlation between VGG-19 network layers and the brain for face and body images in intact, high and low frequencies.

According to the above 2 diagrams, the best layers for most images are the initial layers, but for intact body images, the end layers are more similar, and for blurry body images, the middle layers.

Also, the maximum correlation for each category of images has been shown.

4-1-2 Conclusion

According to the above results, it can be concluded that the similarity between the IT region and the layers of the neural networks for intact and blurry body images are similar to each other, and sharp body images are different from them, but for face images, all three frequencies behave like each other. Also, the final layers of the neural networks are not necessarily more similar to the IT region, and this hypothesis is rejected according to these results.

By examining the highest correlation of different categories, we find that blurry face images have the highest correlation and blurry body images have the lowest correlation.

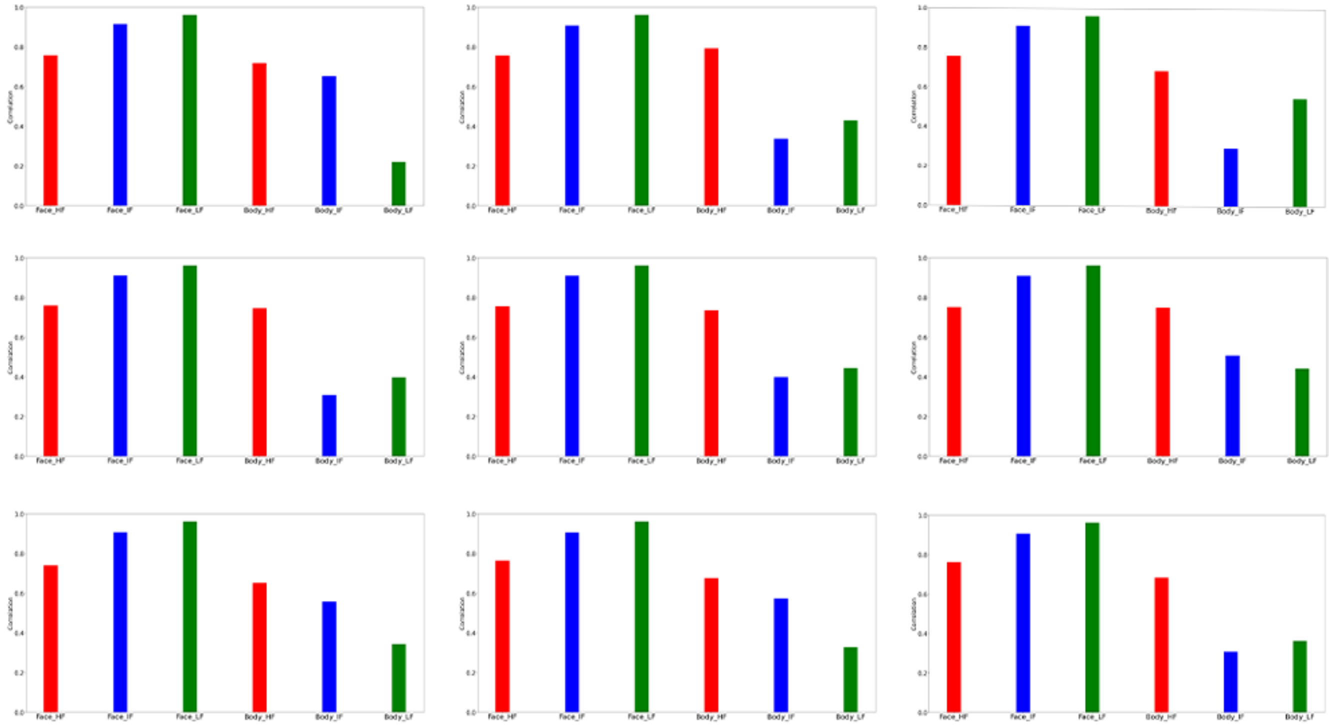

Figure 51: The highest correlation of different categories with the IT region of the brain.

Also, in most categories, the initial layers have the highest correlation with the IT region of the brain, except for intact body images, which have more similarity in the final layers, and blurry body images, which have the highest correlation in the middle layers.

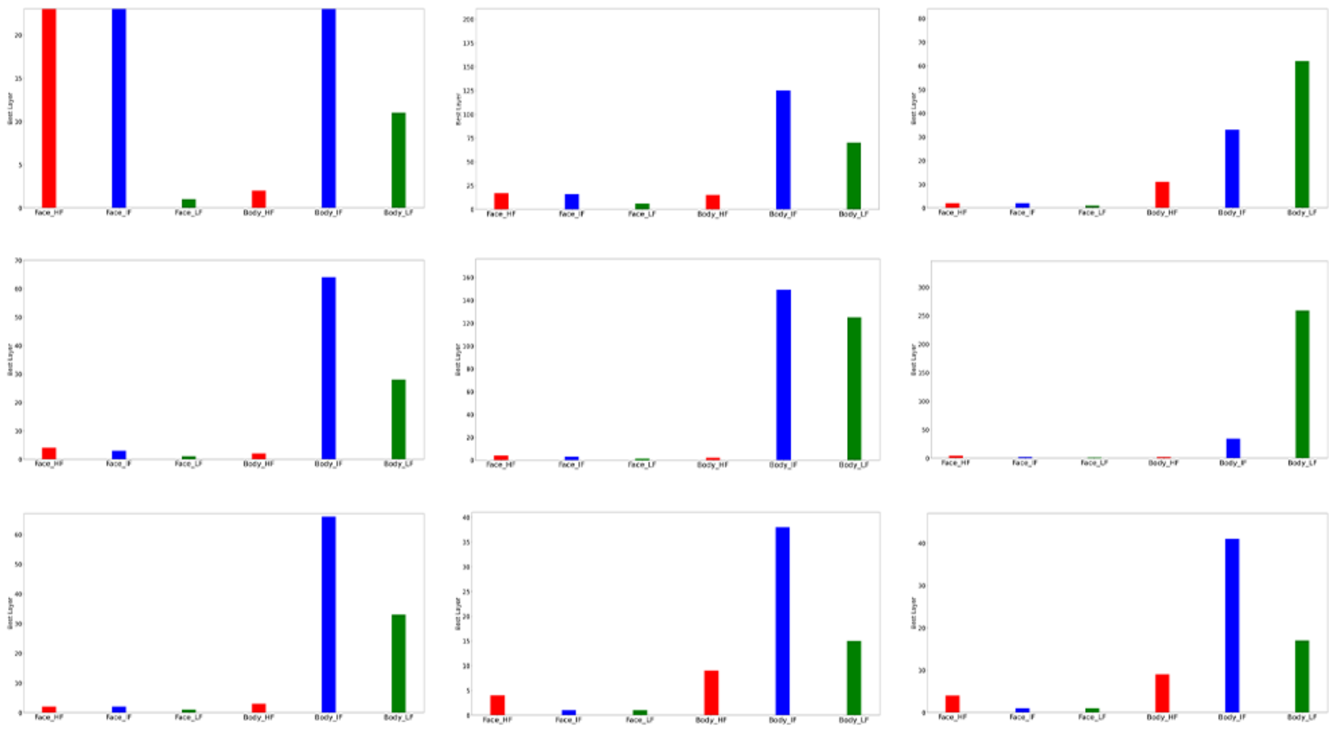

Figure 52: The layer of neural networks with the highest correlation in different categories with the IT region of the brain.

4-2-1 Examining the effect of frequency on neural networks and the brain

In this section, we compare the effect of frequency on the function of the IT region of macaque's brain and the pre-classifier layer of AlexNet, GoogleNet, MobileNet-V2, ResNet18, ResNet50, Resnet101, VGG-16, VGG-19 and SqueezeNet1.1 neural networks. As mentioned, this is done by training a new classifier with images of one set of frequencies and testing it with images of other frequencies. For this, the Boot Strapping method is used to obtain a 95% confidence interval. We plot the average of 200 repetitions of this test and also specify the 95% confidence interval on the graphs.

Here, we examine 6 body images, 3 human bodies and 3 animal bodies, as well as 6 face images, 3 human faces and 3 animal faces, each in 3 intact, blurry and sharp frequencies. 6 face images and 6 body images are selected randomly with placement, and training and testing are done on them every time. So, in total, for each frequency, we have 12 images in two categories of body and face.

First, we do this for the IT region of the brain:

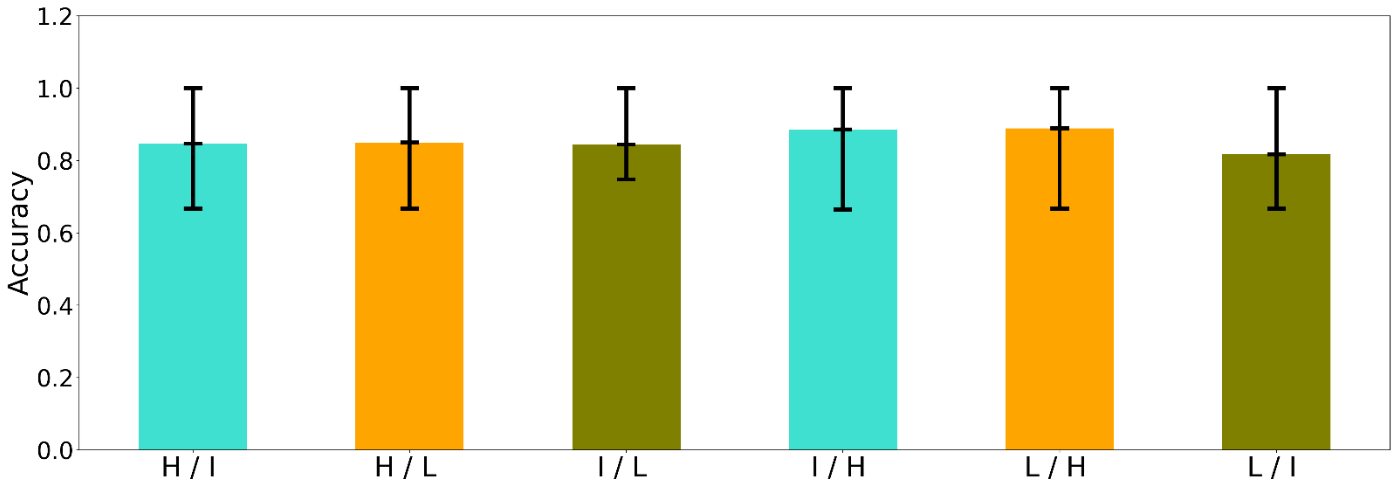

Figure 53: Training a classifier with brain's data. Each time, it is trained with data from one frequency and tested with data from another frequency. H stands for high frequency, or sharp images, L stands for low frequency, or blurry images and I stands for intact frequency, or intact images. For example, H/L means the classifier is trained with sharp image data and tested with blurry images. Finally, the accuracy of this category is averaged in 200 repetitions of this process, and its 95% confidence interval is also drawn.

We can see that the function of the IT cortex of the brain is very uniform and almost unchanged in all frequencies for face and body images. This means that the IT region does not change in performance when the frequency of images changes and is resistant to frequency changes. Also, it has about 80% accuracy in recognizing images of other frequencies, which is a decent performance.

Brain's data was averaged between 70 and 200 milliseconds after showing the images. Now, we check the brain data for 50 millisecond intervals with 5 millisecond strides between the intervals to have more confidence and measure the brain's performance at different moments. The average accuracy of all 6 combinations is plotted with one standard deviation from the average.

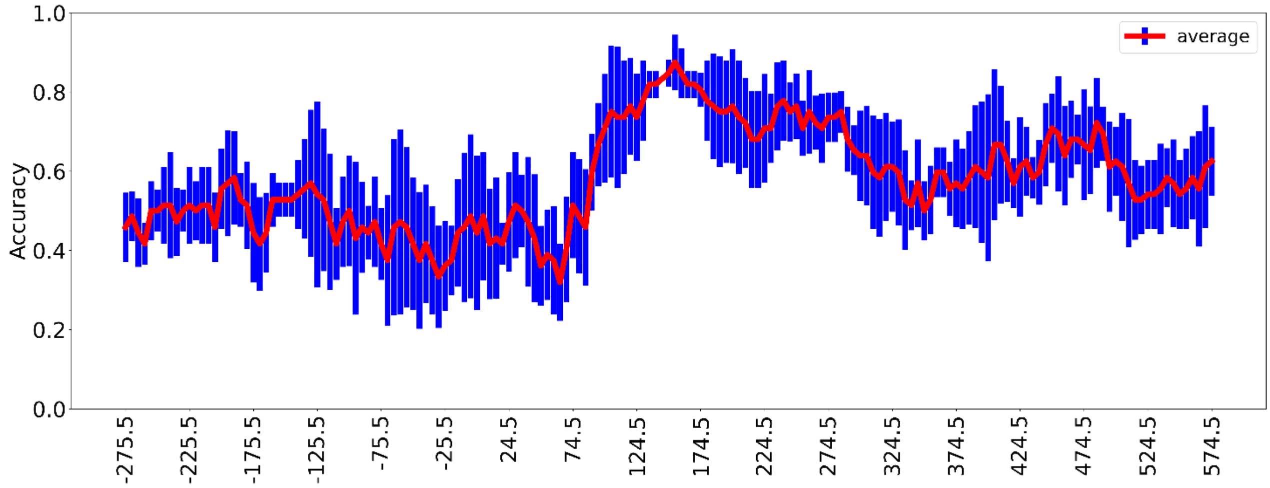

Figure 54: Average accuracy of 6 different modes along with a standard deviation.

As we can see, the accuracy in 70 milliseconds after displaying the image has a significant jump from the 50% accuracy in the intervals before that, and until 200 milliseconds after displaying the images, the accuracy is still high, so our assumption for the interval of 70 to 200 milliseconds was correct.

Now we do the same process for neural networks:

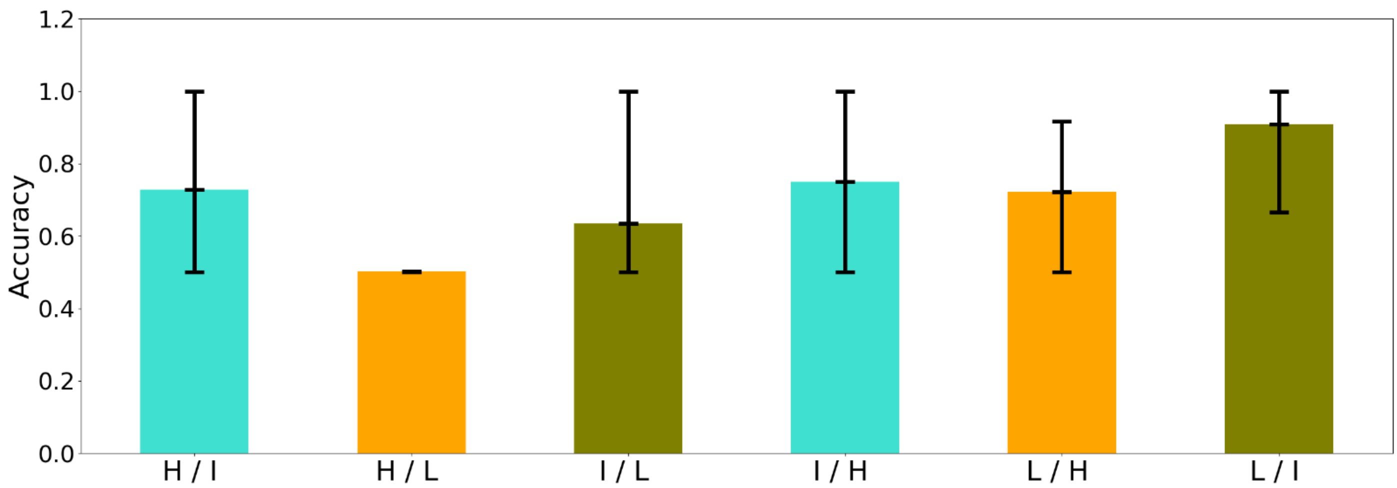

Figure 55: Training a classifier with AlexNet data and testing it with other frequencies.

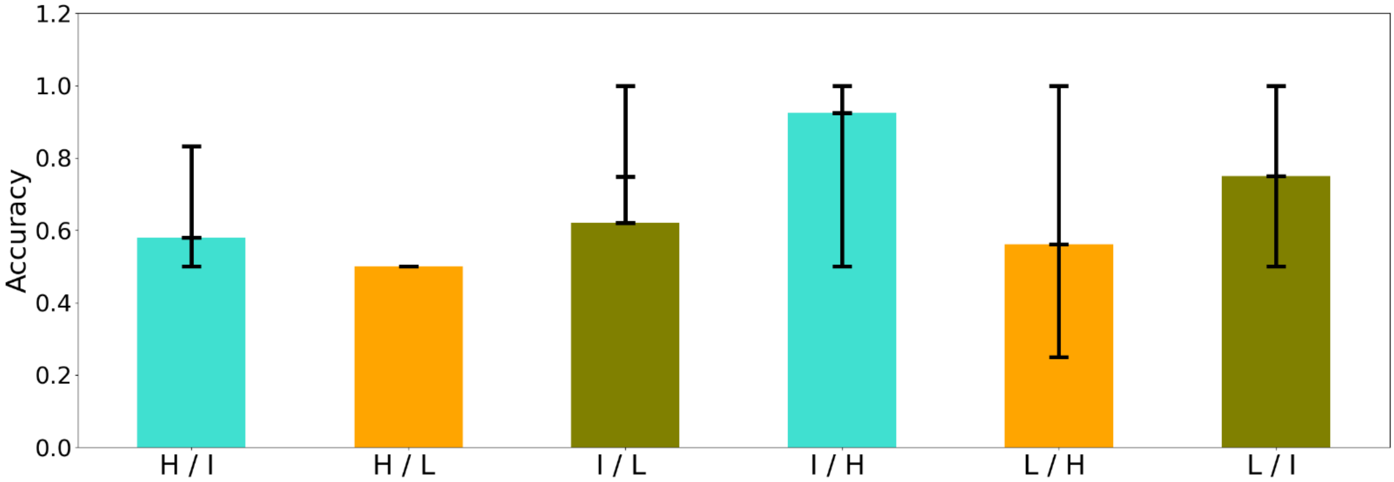

Figure 56: Training a classifier with GoogleNet data and testing it with other frequencies.

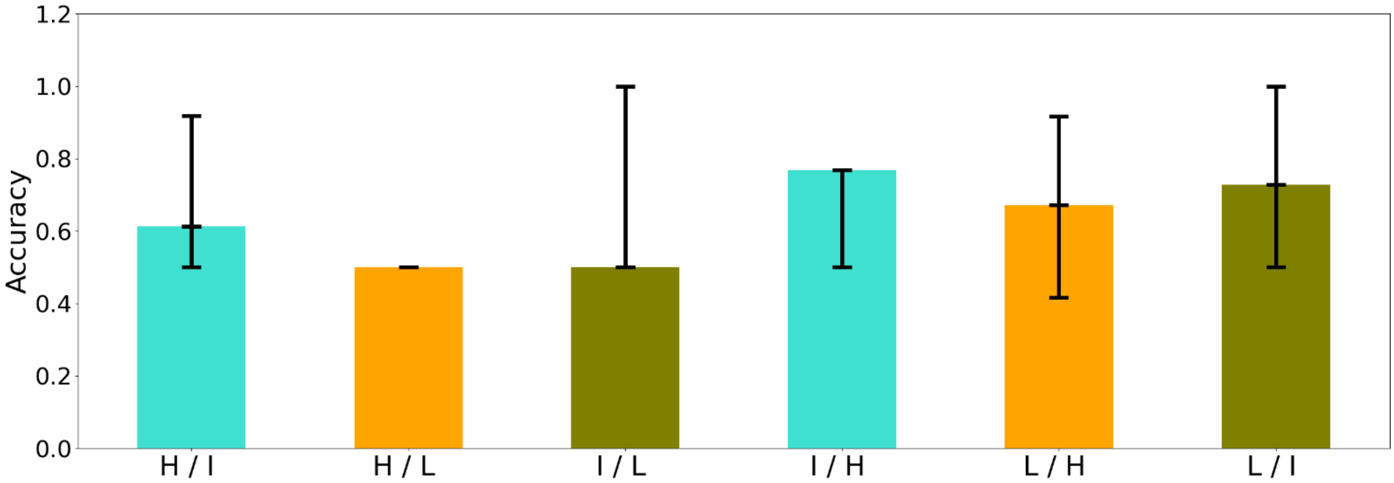

Figure 57: Training a classifier with MobileNet-V2 data and testing it with other frequencies.

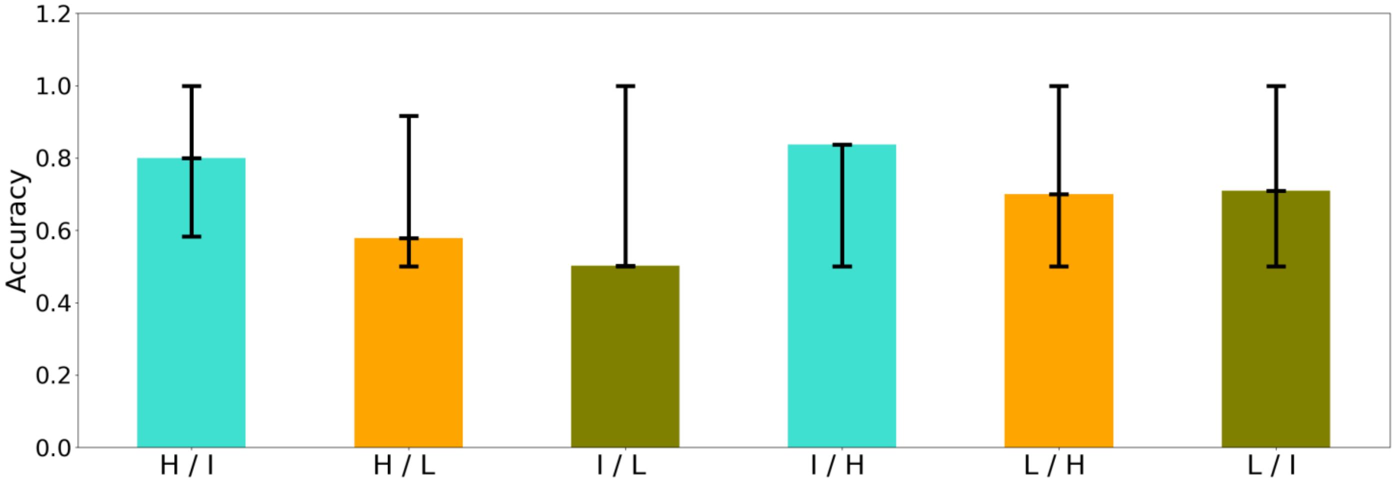

Figure 58: Training a classifier with ResNet18 data and testing it with other frequencies.

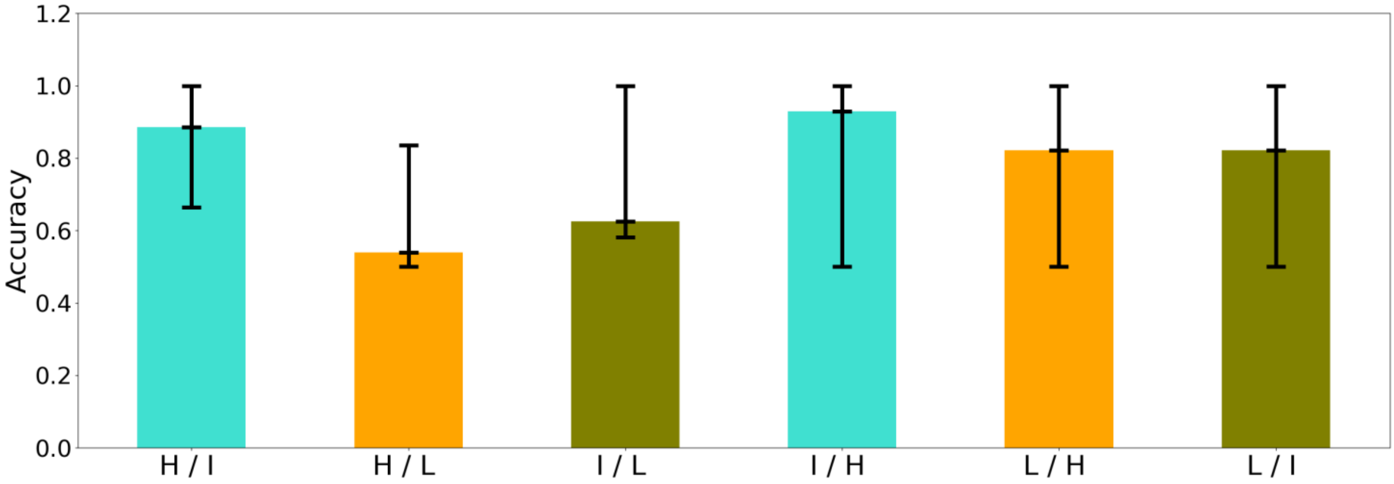

Figure 59: Training a classifier with ResNet50 data and testing it with other frequencies.

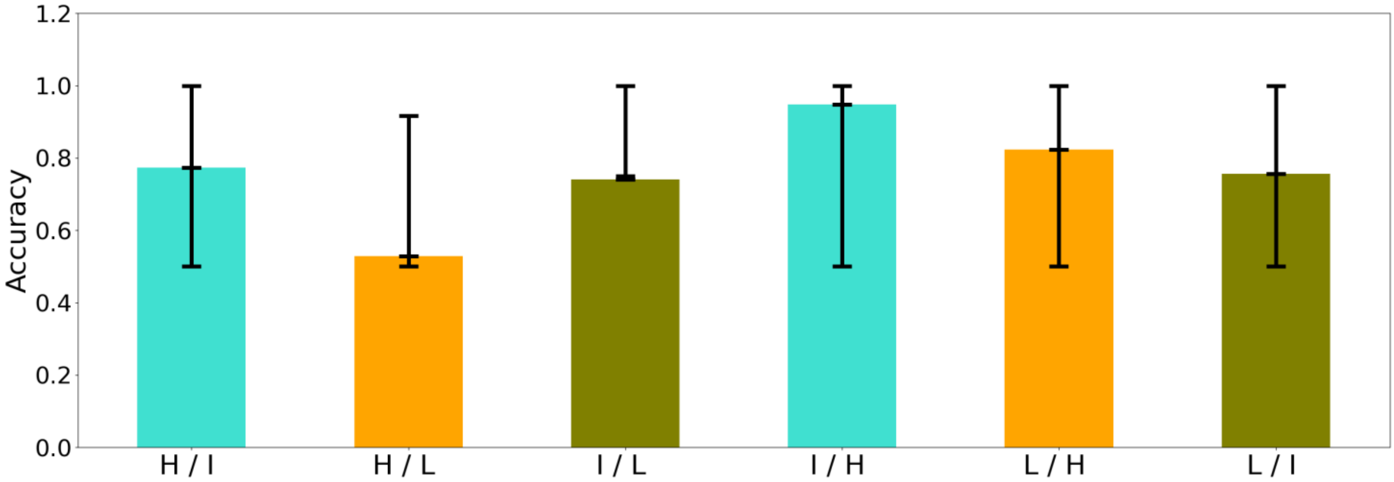

Figure 60: Training a classifier with ResNet101 data and testing it with other frequencies.

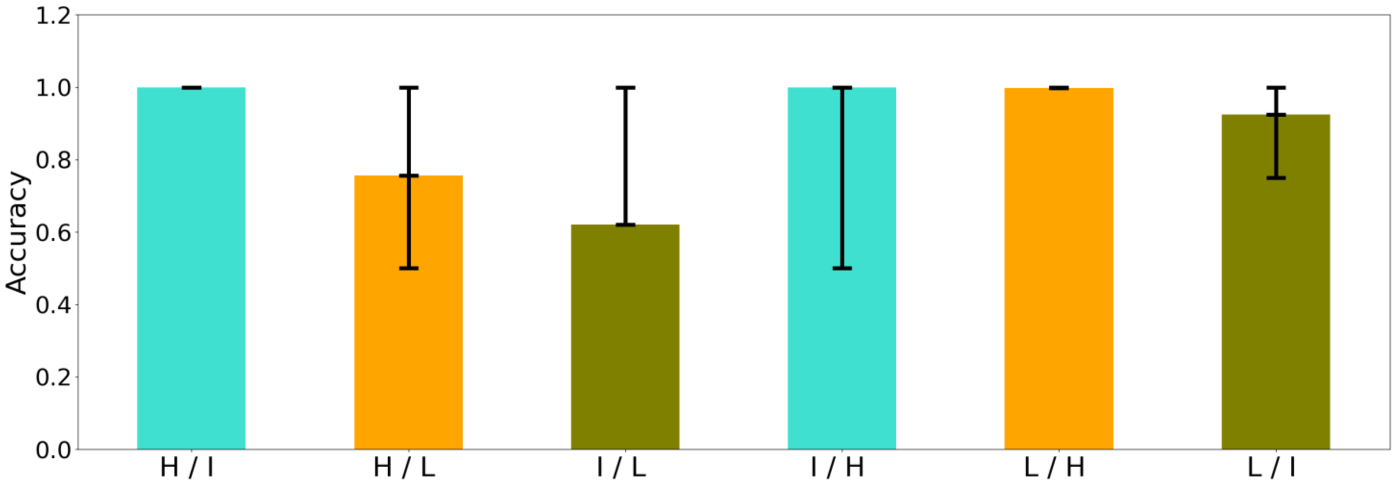

Figure 61: Training a classifier with SqueezeNet1.1 data and testing it with other frequencies.

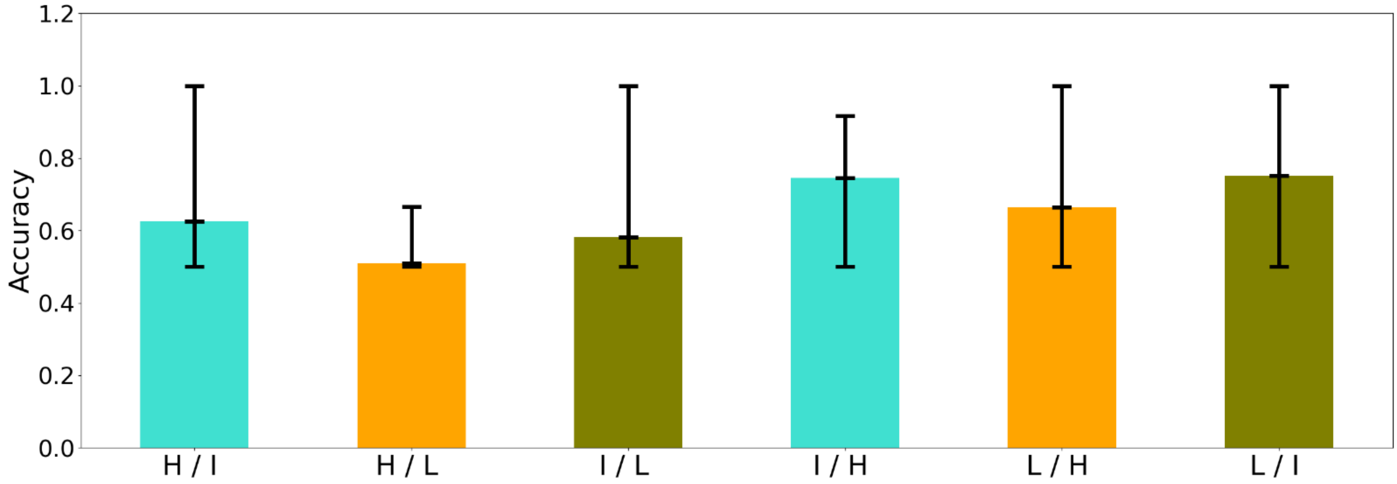

Figure 62: Training a classifier with VGG-16 data and testing it with other frequencies.

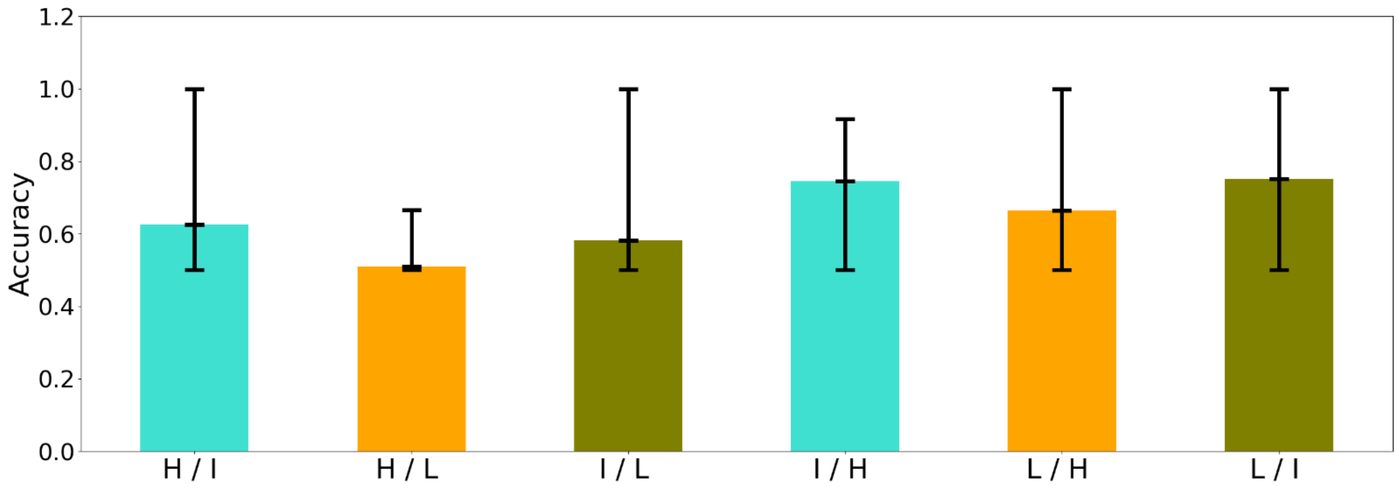

Figure 63: Training a classifier with VGG-19 data and testing it with other frequencies.

As it can be seen, the effect of frequency on neural networks is significant, the frequency of the images with which a neural network is trained has a tremendous effect on the decision of that neural network for images with other frequencies. And this effect is different for different frequencies and different combinations.

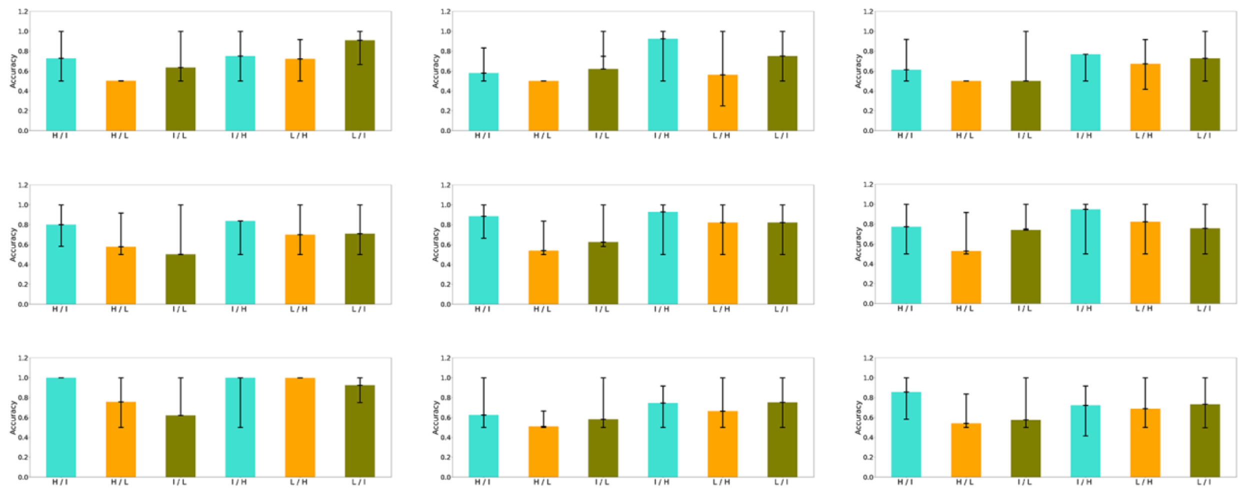

Figure 64: Comparison of the effect of frequency on 9 different neural networks of the CNN type

By calculating the Euclidean distance of the accuracy of each neural network from the brain's accuracies, the best and worst networks in terms of similarity with the brain, are as follows:

1. SqueezeNet1.1
2. ResNet50
3. VGG-19
4. ResNet101
5. MobileNet-V2
6. ResNet18
7. AlexNet
8. VGG-16
9. GoogleNet

According to the results of the effect of image frequency on neural networks, we can conclude the following results:

- Training a neural network with intact images and testing it with sharp images has always a better accuracy than vice versa.

  I/H > H/I

- Training a neural network with blurry images and testing it with intact images always has a better accuracy than vice versa.

  L/I > I/L

- Training a neural network with blurry images and testing it with sharp images always has better accuracy than vice versa.

  L/H > H/L

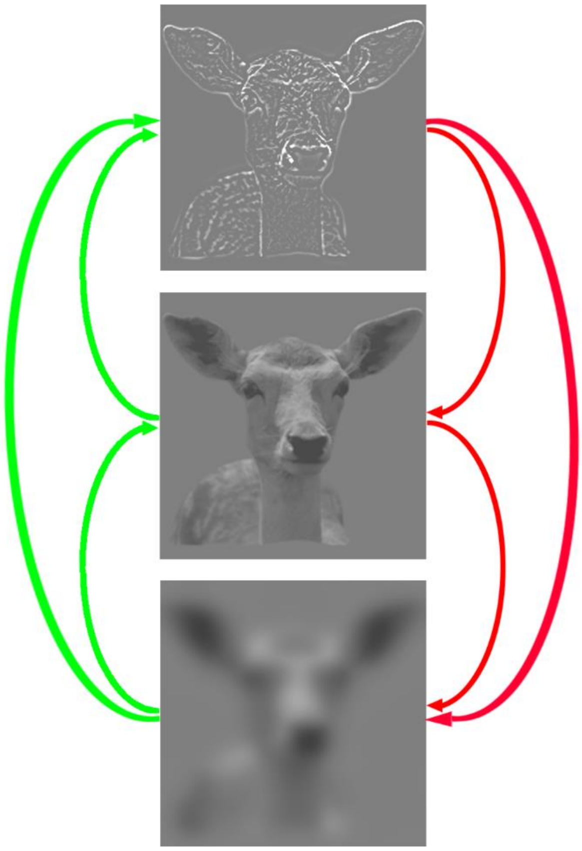

Figure 65: In some combinations of training and testing networks, training with one frequency and testing with another frequency leads to higher accuracy than vice-versa.

4-3 Summary

In this chapter, we examined the similarity of different layers of CNN neural networks with the IT cortex of macaque's brain for face and body images at different frequencies and we came to the conclusion that the final layers of a neural network are not necessarily more comparable with the IT cortex of the brain than the initial layers, so, it cannot be said that the final layers of the brain and neural networks are functionally similar. Also, we examined the effect of frequency in the decision-making of the IT region of the brain and the layer before the classifier layer of neural networks and we saw that the IT region does not change much in terms of performance for body and face images for different frequencies, but this difference in neural networks is very significant, so neural networks are very sensitive to frequency, but the IT cortex is not.

Chapter 5:

Conclusion and Suggestions

5-1 Conclusion

As we have seen, the effect of frequency on neural networks is significant. The frequency of the images with which a neural network is trained affects the decision of that neural network for images with other frequencies and this effect is different for different frequencies. Meanwhile, the IT cortex of macaque's brain is resistant to frequency changes and its accuracy does not change.

Also, we have seen that the claim that the final layers of the brain and the final layers of neural networks are functionally comparable is wrong. In some frequencies, the correlation changes between the brain and the neural network layers are different from other frequencies, for example, the correlation changes for intact and blurry frequencies are almost similar, but the sharp frequency images behave differently. Also, in most cases, the initial layers of the neural networks are most similar to the visual IT cortex of the brain, but in some cases, the final layers are more comparable. In addition, the values of the greatest correlation of layers in body and face images in different frequencies are different.

5-2 Suggestions

According to the results obtained in the previous sections, to design new neural networks, we must keep the following points in mind:

1. Try to design a network whose function in the final layers is similar to the final layers, for example the IT cortex, of the brain. As we have seen, the IT Cortex performs very well between frequency changes and does not suffer from accuracy losses. If this is done, the new neural networks will be resistant to changes in the frequency of the images.

2. The behaviors of the layers of neural networks in blurry and intact frequencies for body and face images compared to the IT cortex of the brain are very similar, yet sharp images have a completely different behavior from the previous two frequencies. It is preferable if new neural networks behave the same in different frequencies compared to the IT region.

REFERENCES

[1] Kriegeskorte, N. Deep neural networks: a new framework for modeling biological vision and brain information processing. Annu. Rev. Vis. Sci. 1, 417–446 (2015).

[2] Rajalingham, R. et al. Large-scale, high-resolution comparison of the core visual object recognition behavior of humans, monkeys, and state-of-the-art deep artificial neural networks. J. Neurosci. 38, 7255–7269 (2018).

[3] Serre, T. Deep learning: the good, the bad, and the ugly. Annu. Rev. Vis. Sci. 5, 21.1–21.28 (2019).

[4] Yamins, D. L. K. & DiCarlo, J. J. Using goal-driven deep learning models to understand sensory cortex. Nat. Neurosci. 19, 356–365 (2016).

[5] Cichy, R. M., Khosla, A., Pantazis, D., Torralba, A. & Oliva, A. Comparison of deep neural networks to spatiotemporal cortical dynamics of human visual object recognition reveals hierarchical correspondence. Sci. Rep. 6, 27755 (2016).

[6] Eickenberg, M., Gramfort, A., Varoquaux, G. & Thirion, B. Seeing it all: convolutional network layers map the function of the human visual system. NeuroImage 152, 184–194 (2017).

[7] Güçlü, U. & van Gerven, M. A. J. Increasingly complex representations of natural movies across the dorsal stream are shared between subjects. NeuroImage 145, 329–336 (2017).

[8] Khaligh-Razavi, S.-M. & Kriegeskorte, N. Deep supervised, but not unsupervised, models may explain IT cortical representation. PLOS Comput. Biol. 10, e1003915 (2014).

[9] Yamins, D. L. K. et al. Performance-optimized hierarchical models predict neural responses in higher visual cortex. Proc. Natl Acad. Sci. USA 111, 8619–8624 (2014).

[10] Cadieu, C. F. et al. Deep neural networks rival the representation of primate IT cortex for core visual object recognition. PLOS Comput. Biol. 10, e1003963 (2014).

[11] Xu, Y., Vaziri-Pashkam, M. Limits to visual representational correspondence between convolutional neural networks and the human brain. Nat Commun 12, 2065 (2021).

[12] Solomon SG and Rosa MGP (2014) A simpler primate brain: the visual system of the marmoset monkey. Front. Neural Circuits 8:96. (2014)

[13] Hung CP, Kreiman G, Poggio T, DiCarlo JJ. Fast readout of object identity from macaque inferior temporal cortex. Science. 2005 Nov 4;310(5749):863-6.

[14] Simo Vanni, Henri Hokkanen, Francesca Werner, Alessandra Angelucci, Anatomy and Physiology of Macaque Visual Cortical Areas V1, V2, and V5/MT: Bases for Biologically Realistic Models, Cerebral Cortex, Volume 30, Issue 6, June 2020, Pages 3483–3517

[15] Brooks, Daniel & Sigurdardottir, Heida & Sheinberg, David. (2014). The neurophysiology of attention and object recognition in visual scenes. 10.13140/2.1.3634.5925.

[16] Gattass R, Nascimento-Silva S, Soares JG, Lima B, Jansen AK, Diogo AC, Farias MF, Botelho MM, Mariani OS, Azzi J, Fiorani M. Cortical visual areas in monkeys: location, topography, connections, columns, plasticity and cortical dynamics. Philos Trans R Soc Lond B Biol Sci. 2005 Apr

[17] https://miro.medium.com/max/1400/1*Gh5PS4R_A5drl5ebd_gNrg@2x.png

[18] https://www.researchgate.net/profile/Jose-Benitez-Andrades/publication/339447623/figure/fig2/AS:862056077082627@1582541593714/A-vanilla-Convolutional-Neural-Network-CNN-representation.png

[19] https://www.telusinternational.com/articles/difference-between-cnn-and-rnn

[20] http://www.ipm.ac.ir/personalinfo.jsp?PeopleCode=IP1300047