CSAI

Assignment-4

Report

Akshit Sharma

(2021101029)

Subject chosen: Subject 1, **fMRI Data File chosen:** CSI1_ROIs_TR34.mat

Data Preprocessing: First, I loaded the data from CSI1_ROIs_TR34.mat file. Then, I separated the left and right hemisphere data for each ROI. This is followed by concatenating the LH and RH data for each ROI separately to get fMRI data corresponding to each ROI. Then, I loaded the stimulus image names in the order of presentation to subject, from the CSI01_stim_lists.txt file. After doing this, I loaded all the unique images from the COCO, ImageNet and Scene folders in the Scene_Stimuli folder. Then, I used pre-trained ResNet 15 model to obtain (512,) dimensional representations for each stimulus image. This will be used as input to encoder and as the ground truth for decoder for each ROI.

Performance Metrics for Encoder and Decoder trained for each ROI

Early Visual

```
Encoder:
Model: Ridge Regressor
Test Mean Squared Error: 0.8875507350234483
Average Pearson Correlation: 0.33044699009038286
R2_score:0.014597538468767157
Decoder:
Model: Ridge Regressor
Test Mean Squared Error: 0.7993132744922933
Average Cosine Similarity: 0.7288274048123661
Average Euclidean Distance: 20.005660847739982
Average Pearson Correlation: 0.20428625861496666
```

LOC (Lateral Occipital Complex)

```
Encoder:
Model: Ridge Regressor
Test Mean Squared Error: 0.8952314672415165
Average Pearson Correlation: 0.32238443725670357
R2_score:0.030097811511646987
Decoder:
Model: Ridge Regressor
Test Mean Squared Error: 0.7885077811394112
Average Cosine Similarity: 0.7328931669166819
Average Euclidean Distance: 19.8707948202259
Average Pearson Correlation: 0.23007889414151486
```

OPA (Occipital Place Area)

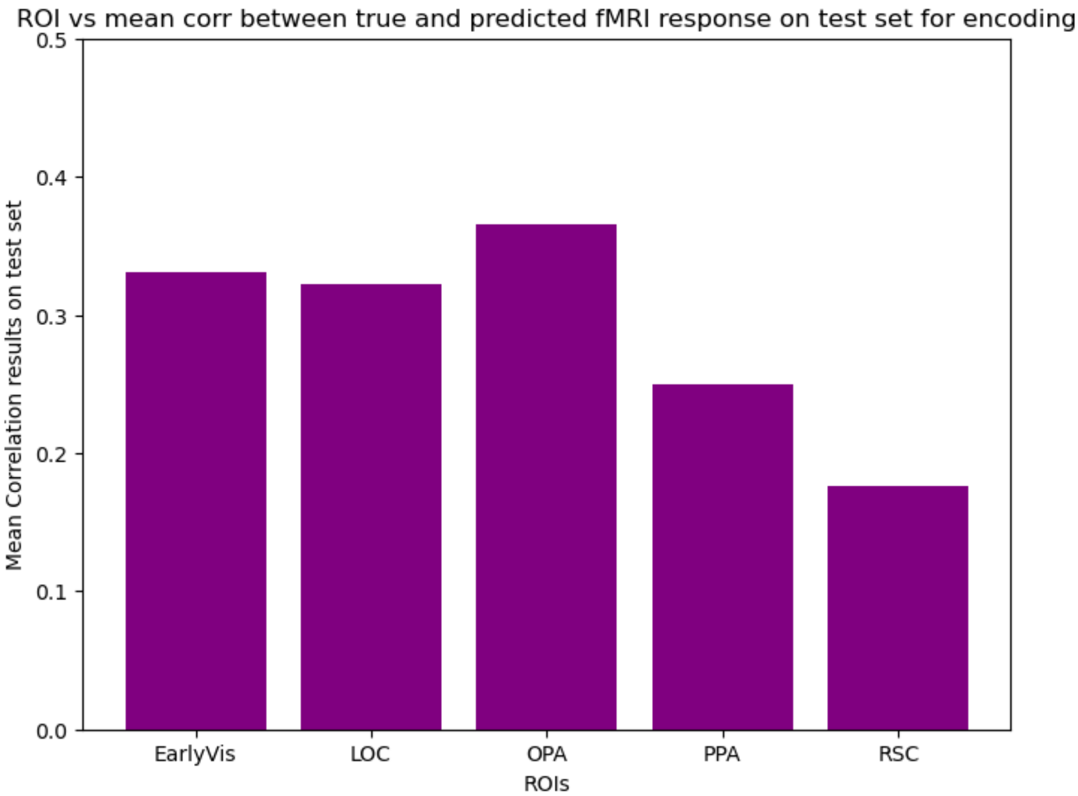
```
Encoder:
Model: Ridge Regressor
Test Mean Squared Error: 0.8618060529772467
Average Pearson Correlation: 0.36556049519888895
R2_score:0.03561457000714202
Decoder:
Model: Ridge Regressor
Test Mean Squared Error: 0.7883215893615161
Average Cosine Similarity: 0.7332713589424891
Average Euclidean Distance: 19.86485362768712
Average Pearson Correlation: 0.23078923036022406
```

PPA (Parahippocampal Place Area)

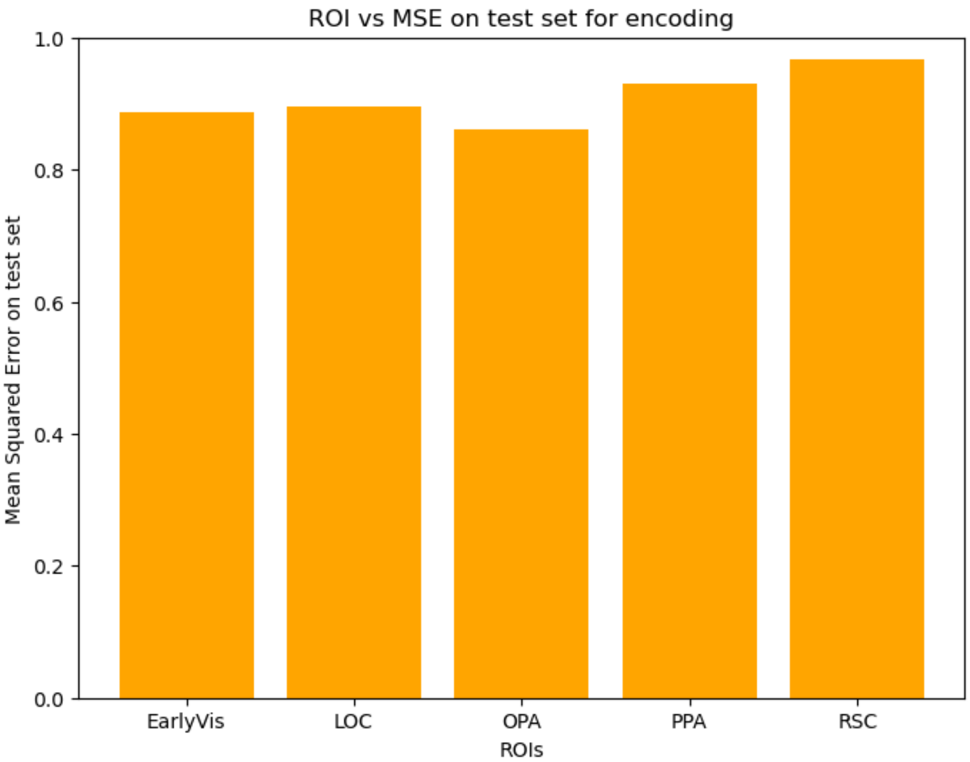
```
Encoder:
Model: Ridge Regressor
Test Mean Squared Error: 0.9302014942437175
Average Pearson Correlation: 0.24934237529626463
R2_score:0.018517890649996316
Decoder:
Model: Ridge Regressor
Test Mean Squared Error: 0.7956247409819694
Average Cosine Similarity: 0.7301576187908684
Average Euclidean Distance: 19.963923499947175
Average Pearson Correlation: 0.21334973757226403
```

RSC (Retrospienial Complex)

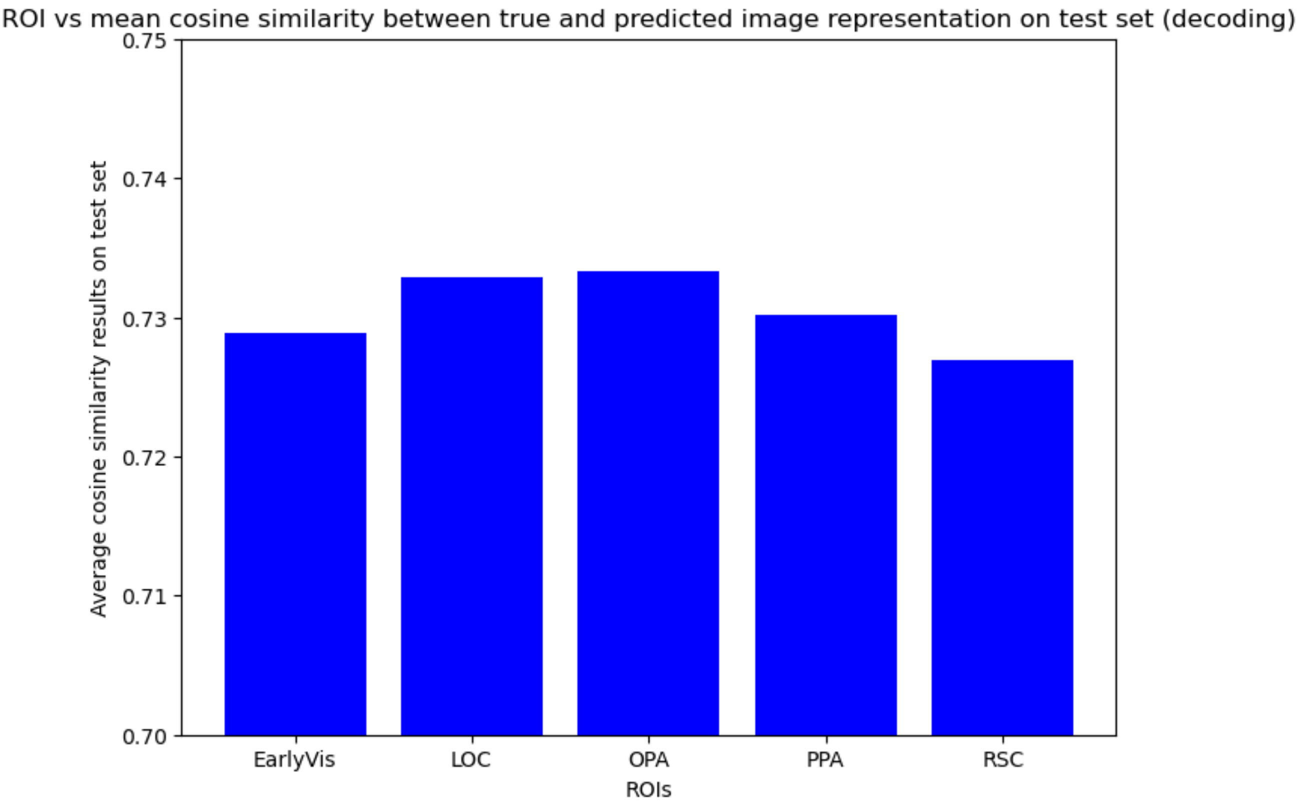
```
Encoder:
Model: Ridge Regressor
Test Mean Squared Error: 0.9662245693231641
Average Pearson Correlation: 0.17584603573507254
R2_score:0.014323425256630658
Decoder:
Model: Ridge Regressor
Test Mean Squared Error: 0.803555721633685
Average Cosine Similarity: 0.7269255276926404
Average Euclidean Distance: 20.064755410365777
Average Pearson Correlation: 0.19010964858439555
```



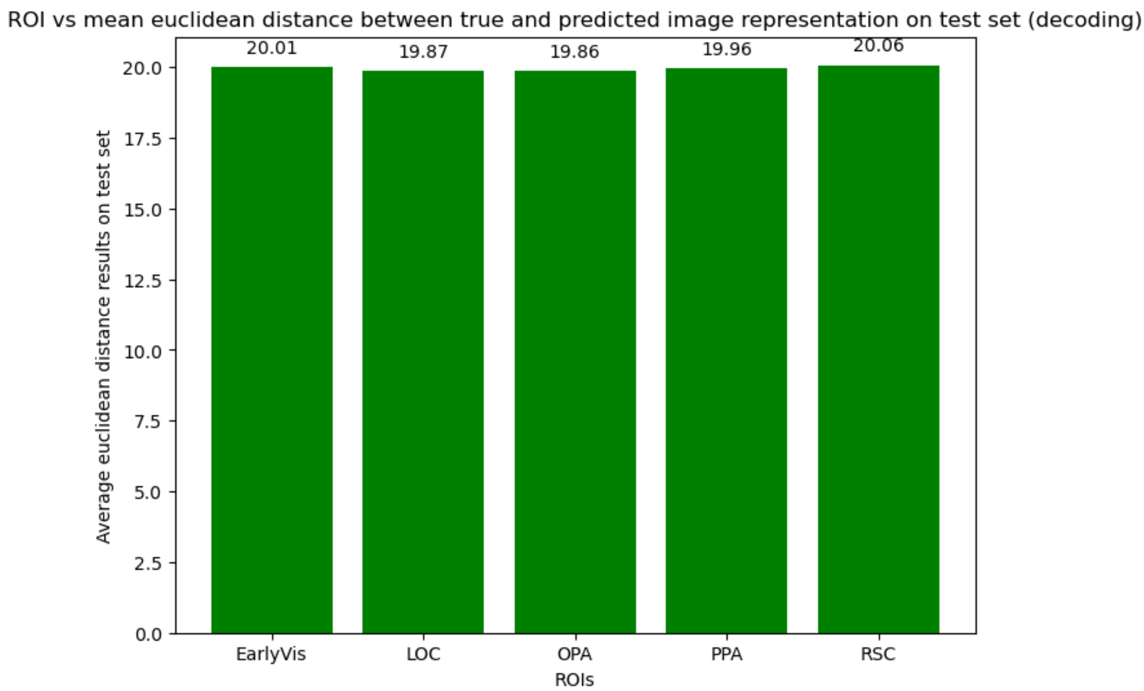
The bar graph above shows that encoder for OPA gives highest performance, shown by the highest value of mean correlation between true and predicted fMRI response on test set.



The bar graph above shows that MSE on test set is least for OPA encoder.



The above bar graph shows that the decoder for OPA gives the highest mean cosine similarity between true and predicted image representations for test data. So, it performs the best among the given ROIs.



The above bar graph shows that the mean euclidean distance between true and predicted image representations is least for OPA decoder (best performing).

Analysis of Observations

Different brain regions might exhibit varying levels of predictability based on the features extracted from ResNet and the nature of the task. For example, early visual areas might be more predictable from visual stimuli, while higher-order regions like PPA or RSC might require more complex feature representations.

Why Ridge Regression?

Ridge regression adds a penalty term to the loss function, which penalises large coefficients. This regularisation helps prevent overfitting, which can be especially important when dealing with high-dimensional data such as fMRI data, where the number of features (voxels or ROIs) exceeds the number of samples. By controlling model complexity, Ridge regression can generalise well to unseen data, leading to better performance on the test set across all ROIs. Ridge regression is less sensitive to outliers compared to ordinary least squares regression, as it penalises large coefficients. Thus, all these features make it suitable for our task.

Which Encoder Model performs best and why?

From the performance metrics, we observe that encoder model for OPA performs the best (gives highest average correlation on test set and least value of MSE). This can be explained as follows:

- OPA is known to be specialised for processing scenes and environments. Its neural activity is particularly sensitive to spatial layouts, landmarks and contextual information within visual scenes.
- Achieving the best encoding performance for the OPA indicates that the features extracted from the ResNet image representations effectively capture the spatial and contextual information relevant to scene processing. This suggests that the OPA may rely heavily on high level visual features extracted by the ResNet for encoding scene related information.
- Scenes are complex visual stimuli that contain a wealth of spatial and contextual information. Compared to simpler stimuli like basic shapes or objects, encoding scene-related information might require more complex and higher-dimensional representations. The ResNet image representations, which are derived from deep convolutional neural networks trained on large-scale visual recognition tasks, are well suited for capturing the richness and complexity of scenes, making them particularly effective for encoding OPA activity.

The second best performance was shown by EarlyVis encoder. Its primary function is to receive and process basic visual input from the eyes, such as detecting edges, orientation, and motion. EarlyVis is crucial for initial processing of visual stimuli and forms the foundation for higher level visual processing in subsequent visual areas. The ResNet features extracted from the last layer of the network may preserve important low level visual information that is highly informative for encoding neural responses in the Early Visual Cortex. The network might have learned to extract features that are particularly relevant for early visual processing, leading to strong encoding performance in EarlyVis.

Decoder performance analysis

We observe that OPA also performs best among the decoding models for all ROIs (based on mean cosine similarity score on test data, average euclidean distance between predicted and true representations and highest mean correlation score for test set). It suggests that this region plays a critical role in both representing and decoding visual information related to scenes. This consistency underscores the importance of the OPA in processing scene related visual stimuli and suggests that its neural activity contains rich information that can be decoded to reconstruct visual representations.

The second best decoding performance is shown by LOC. It integrates information from the Early Visual Cortex and other visual areas to form representations of objects, their shapes and their spatial relationships. LOC is implicated in various visual tasks like object recognition, scene perception and visual attention. The LOC contains neural representations that are highly discriminative for encoding object related visual information. These representations may capture detailed information about object shapes, contours, and spatial arrangements, which are essential for reconstructing image representations from fMRI data. The ResNet image representations, coupled with the distinctive neural activity patterns in the LOC, enable better decoding of object related visual information compared to other ROIs.

<u>RSC give worst encoding as well as decoding performance</u>

The Retrosplenial Cortex is implicated in spatial navigation, memory, and spatial awareness. It integrates spatial information from multiple sensory modalities and plays a role in updating and representing one's spatial orientation in the environment.

The stimuli used in the encoding and decoding tasks does not engage the spatial navigation or memory related functions of RSC. If the stimuli primarily consist of visual scenes, objects, or shapes that do not require spatial navigation or memory processing, the RSC may not exhibit strong neural responses, resulting in poor encoding and decoding performance. The RSC's specialisation for spatial cognition may limit its ability to effectively encode or decode visual information unrelated to spatial tasks.

RSC may have limited sensitivity to the visual features extracted by the ResNet image representations. Its neural activity may be driven more strongly by spatial and contextual information rather than low level visual features. Consequently, the ResNet image representations may not effectively capture the information necessary to encode or decode neural responses in the RSC, leading to poor performance in both tasks.