Effect of Participant Variation on Visual Stimuli Reconstruction Performance From fMRI Signals Using Machine Learning

Author: Quan Zheng, q.zheng-7@student.tudelft.nl
Supervisor: Xucong Zhang

TU Delft

1. Background

Reconstruction of visual stimuli from neuronal activity: Decoding the human brain's activation patterns to re-create images shown to a subject as accurately as possible.

Applications in brain machine interfaces (e.g. Neuralink) and researching dreams.

It us unclear what effect participant variation has on the performance of reconstruction models.

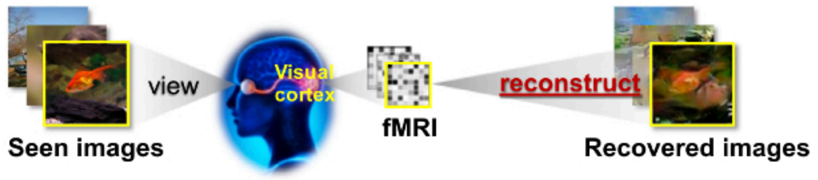

Figure 1: Visual stimulus reconstruction pipeline [1]

2. Research Questions

Is there a predictive ability of participant selection on image reconstruction from fMRI signals using machine learning?

Sub-questions:

- Does any one participant's fMRI data consistently result in better reconstruction performance?
- Is there a significant difference in reconstruction performance between participants?
- What pros and cons do differing image similarity metrics bring to the table?

No two individuals have the same brain anatomy [2]. Our research can indicate the generalizability of machine learning models to neural firing patterns of differing individuals and give guidance on trade-offs between more participants while taking fewer fMRI scans (higher chance of suited participants), or vice versa.

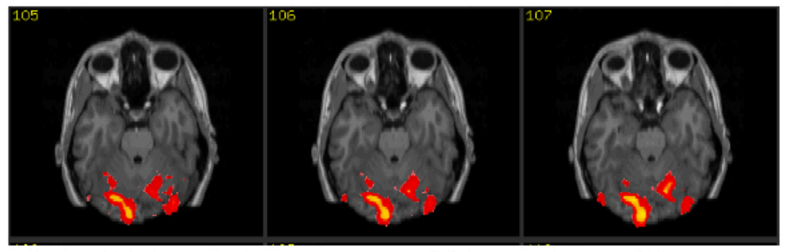

Figure 2: Temporal sequence of fMRI recordings. [3]

3. Methodology

Reconstruction of images was conducted using the Self-Supervised Image Reconstruction machine learning architecture proposed by Gaziv et al. [4].

The approach employs a neural-network based encoder decoder architecture, where an encoder E is trained to predict the fMRI responses to a visual stimulus, with a decoder D being trained to reconstruct the original stimulus from said responses. This architecture enables training using images that are "unpaired" (i.e. with no fMRI response present)

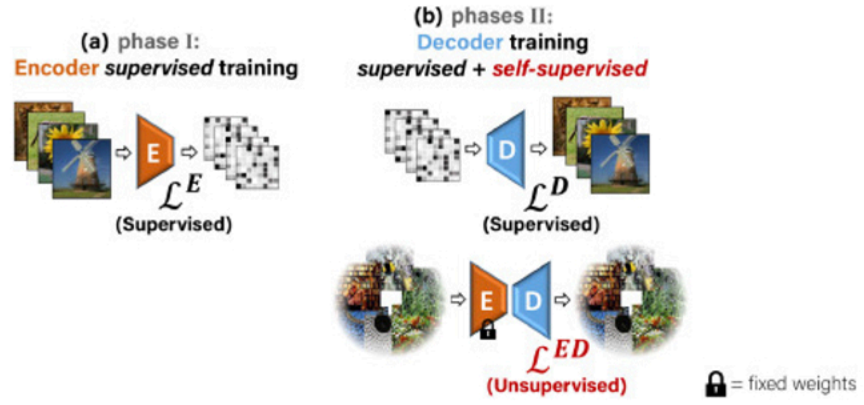

Figure 3: Encoder Decoder architecture in two phases. [4]
a) Phase I trains an encoder to map images to fMRI
b) Phase II trains a decoder to map fMRI to RGB images
This approach allows for self-supervised training using unpaired images

Image reconstruction performance was evaluated using three pixel-wise, and two structural image similarity evaluation metrics. Statistical analysis was conducted by aggregating results by subject (mean, std, min, max, # best reconstructions) and Analysis Of Variance (ANOVA).

Pixel-wise metrics:
- Root Mean Squared Error (RMSE)
- Peak Signal to Noise Ratio (PSNR)
- Signal to Reconstruction Error ratio (SRE)

Structural
- Structural Similarity Index (SSIM)
- Feature-based Similarity Index (FSIM)

Similarity measure	Score
RMSE	0.01064
PSNR	39.3884
SRE	38.0829
SSIM	0.82997
FSIM	0.34981

Similarity measure	Score
RMSE	0.01820
PSNR	34.79547
SRE	42.46683
SSIM	0.80673
FSIM	0.40393

Figure 4: Example reconstructions with corresponding similarity scores

4. Results

Score aggregations by subject can be seen below. Best reconstruction results for pixel-wise evaluation consistently come form subject 3's data, followed by subject 2.

Table 1: RMSE Aggregations by Subject

sub	mean	std	min	max	# best
sub1	0.0202	0.0059	0.0090	0.0373	7
sub2	0.0197	0.0059	0.0091	0.0353	12
sub3	0.0187	0.0055	0.0075	0.0324	18
sub4	0.0205	0.0069	0.0096	0.0437	7
sub5	0.0199	0.0045	0.0105	0.0304	6

Table 2: PSNR Aggregations by Subject

sub	mean	std	min	max	# best
sub1	34.1764	2.5043	28.5723	39.9585	7
sub2	34.3978	2.6221	29.0467	40.7438	12
sub3	34.8819	2.6194	29.7855	42.2375	18
sub4	34.1649	2.7694	27.1911	40.2398	7
sub5	34.1855	2.0530	30.1917	39.5482	6

Table 3: SRE Aggregations by Subject

sub	mean	std	min	max	# best
sub1	42.0562	3.8313	32.2476	50.9959	8
sub2	42.1409	3.7663	33.5379	51.9050	10
sub3	42.3985	3.8109	33.4778	50.6304	18
sub4	42.0798	3.8849	32.0999	50.0271	8
sub5	42.0801	3.5543	34.7736	49.6416	6

Reconstruction performance as measured by structural similarity also consistently favors subject 3, but is inconsistent when ranking remaining subjects.

Table 4: SSIM Aggregations by Subject

sub	mean	std	min	max	# best
sub1	0.7546	0.1110	0.3930	0.9362	9
sub2	0.7553	0.1075	0.3932	0.9555	9
sub3	0.7721	0.1000	0.4117	0.9571	18
sub4	0.7493	0.1149	0.3295	0.9366	7
sub5	0.7413	0.1004	0.4717	0.9309	7

Table 5: FSIM Aggregations by Subject

sub	mean	std	min	max	# best
sub1	0.34703	0.02990	0.27298	0.42097	7
sub2	0.34867	0.02716	0.30908	0.44404	4
sub3	0.36255	0.02694	0.28770	0.44948	21
sub4	0.35757	0.03425	0.28244	0.47116	14
sub5	0.34534	0.02392	0.30092	0.42178	4

ANOVA analysis with $p < 0.05$ yields significant results for performance evaluation using FSIM.

Table 6: ANOVA statistics by performance measure

metric	F-value	p-value
RMSE	0.66159	0.61924
PSNR	0.73673	0.56770
SRE	0.07071	0.99084
SSIM	0.55854	0.69298
FSIM	3.38634	0.01017

5. Discussion

- RMSE, PSNR, and SRE rank subject 3 as best, with similar ranking for remaining participant ranking as pixel-wise evaluation metrics.
- Structural metrics give a more wholistic view of similarity, also favoring subject 3, with inconsistencies ranking other subjects.
- Subject 3 consistently performed best on all five evaluation-metrics, with at least 50% more best performances in every case.
- ANOVA shows significant variance between scores by subject for evaluation using FSIM ($p = 0.01$), with others being non-significant.

6. Conclusion

Reconstructions using subject 3's consistently perform best across the evaluation metrics employed in our paper. This points to participant selection playing a role for reconstruction performance.

ANOVA yields significant variance for performance between subjects as measured using FSIM. Thus, FSIM may have a higher sensitivity to changes from the original image than the remaining metrics. This should be further investigated as our sample size of reconstructed images is limited.

7. Future Work

The Sample size of reconstructed images is limited (5 subjects * 50 reconstructions each). Future work should investigate results found for FSIM further using a larger sample size data set to strengthen the findings of our paper.

Differing models may favor differing brain anatomies or have varying susceptibility to noise. Future research should examine weather the results found in this paper generalize to other reconstruction models.

Acknowledgements

We would like give special thanks to Xucong Zhang for acting as our supervisor and responsible professor for this project, as well as Nikita Smolin for running model training and reconstruction on their personal computer.

References

[1] Ziqi Ren, Jie Li, Xuetong Xue, Xin Li, Fan Yang, Zhicheng Jiao, and Xinbo Gao. Reconstructing seen image from brain activity by visually-guided cognitive representation and adversarial learning. NeuroImage, 228:117602, 2021.

[2] Seyed Abolfazl Valizadeh, Franziskus Liem, Susan Merillat, Jürgen Hänggi, and Lutz Jäncke. Identification of individual subjects on the basis of their brain anatomical features. Scientific Reports, 8(1), 2018.

[3] FMRI functional Magnetic Resonance Imaging Lab, https://home.csulb.edu/-cwallis/482/fmri/fmri.html (accessed Jun. 27, 2023).

[4] Guy Gaziv, Roman Beliy, Niv Granot, Assaf Hoogi, Francesca Strappini, Tal Golan, and Michal Irani. Self-supervised natural image reconstruction and large-scale semantic classification from brain activity. NeuroImage, 254:119121, 2022.7