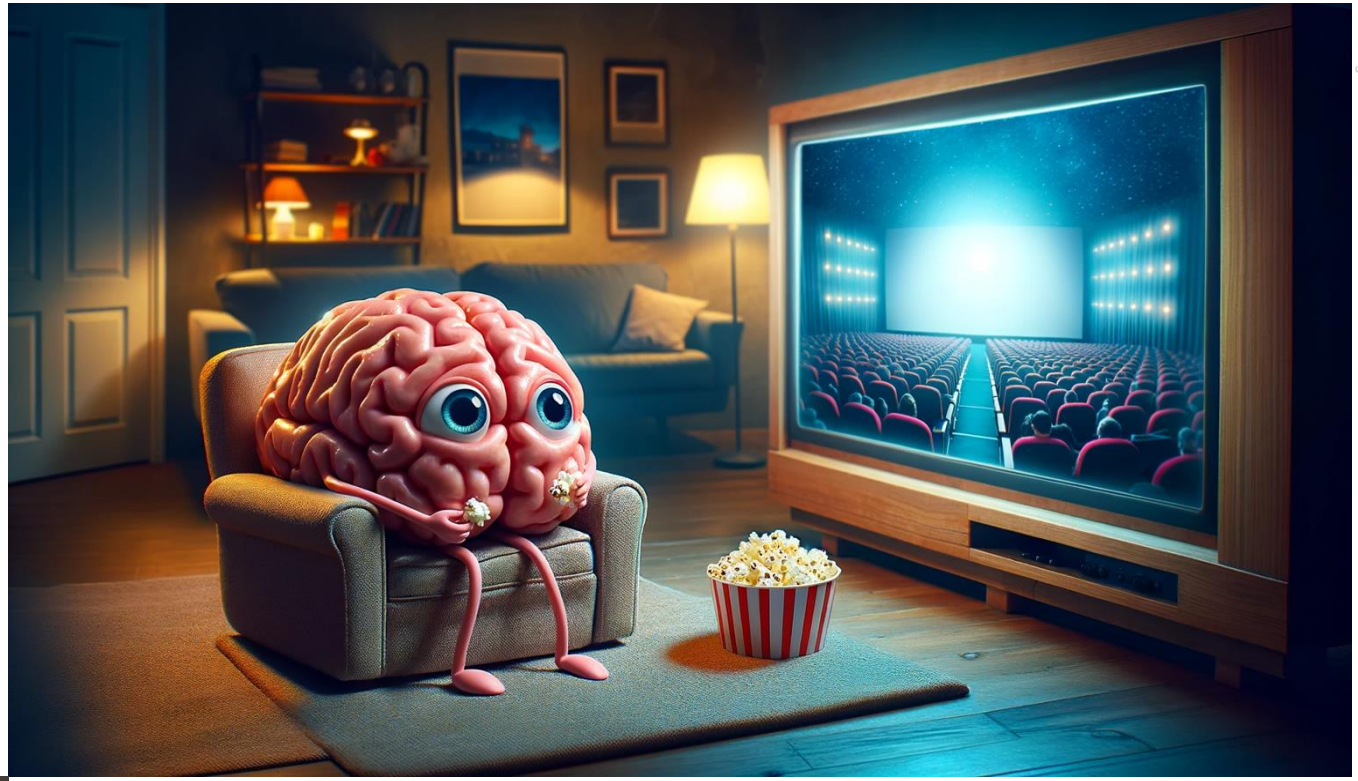
From movie frames to the brain...



This image was created with the assistance of DALL·E

... and back!

Florian David
Neuro-X student
Supervised by
Michael Chan

June 11th, 2024

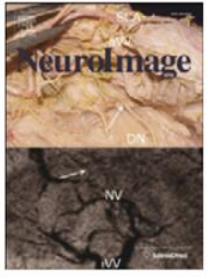
Literature survey

ELSEVIER

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

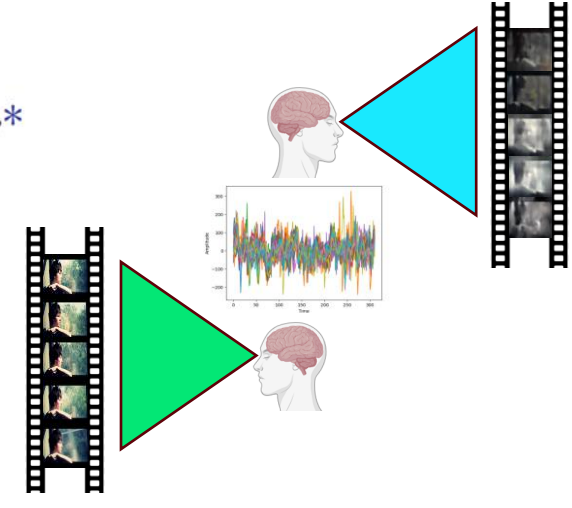


Review

Encoding and decoding in fMRI

Thomas Naselaris [a], Kendrick N. Kay [b], Shinji Nishimoto [a], Jack L. Gallant [a,b,*]

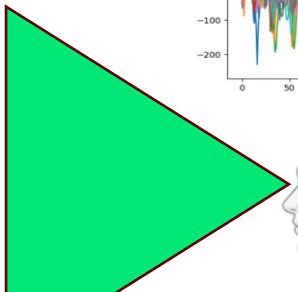
[a] *Helen Wills Neuroscience Institute, University of California, Berkeley, CA 94720, USA*
[b] *Department of Psychology, University of California, Berkeley, CA 94720, USA*

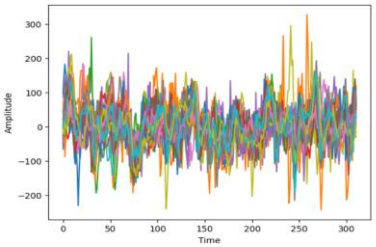


Scientific aim(s)

1.

Build an **encoder** to predict brain activity while someone is watching a movie





2.

Build an **decoder** to validate the predicted brain activity

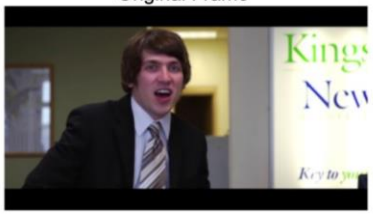
3.

Merge both models into an end-to-end **encoder-decoder** to generate fMRI data on unseen movies

Dataset

Preprocessing on all 14 movies:
- 112x112 pixels
- 32 frames per TR

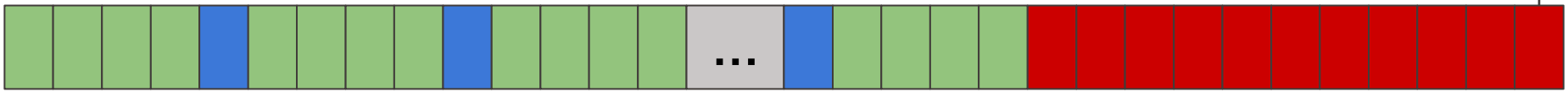
Original Frame



Preprocessed Frame

1 TR = 32 frames

Splitting of 13/14 movies



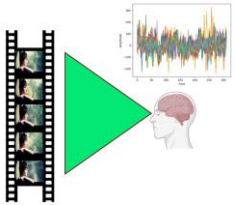
Train Set ~ 64%

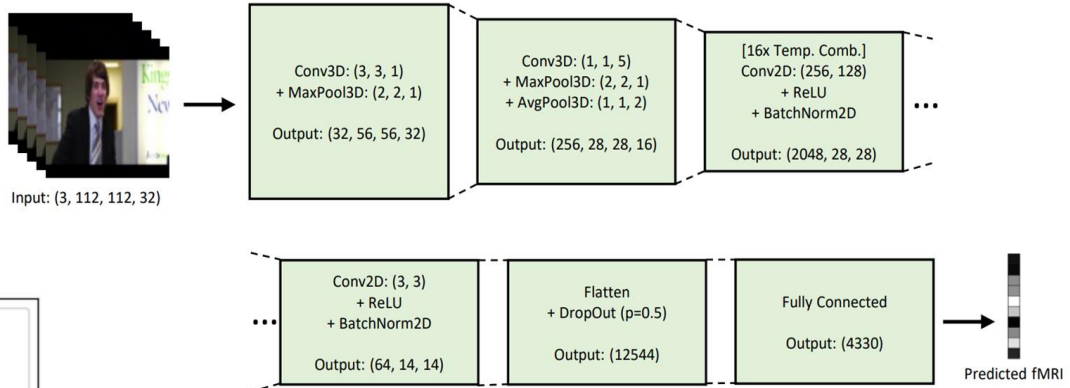
Validation Set ~ 16%

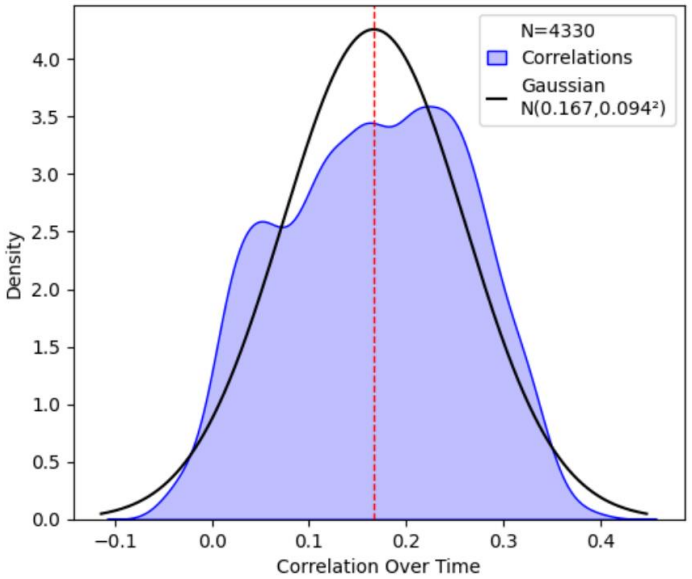
Test Set ~ 20%

1 full movie in the Test Set:
You Again (613TR)

Encoder 1.0







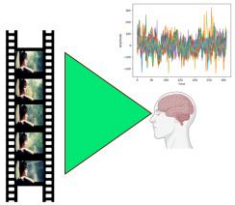
⇒ Best results on the **validation set**

ML4Science results

→ Average correlations on the **validation set**: 0.167
→ Average correlations on the **test set**: 0.001

Can we improve it?

Encoder 2.0

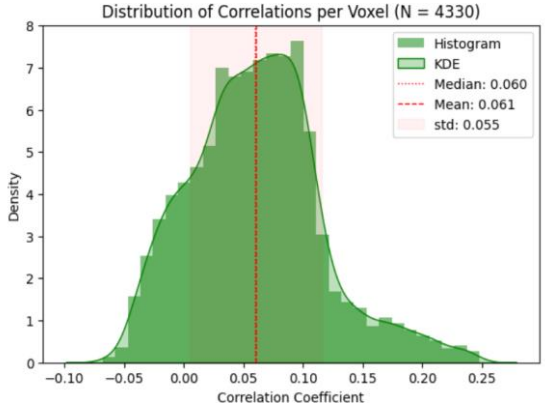


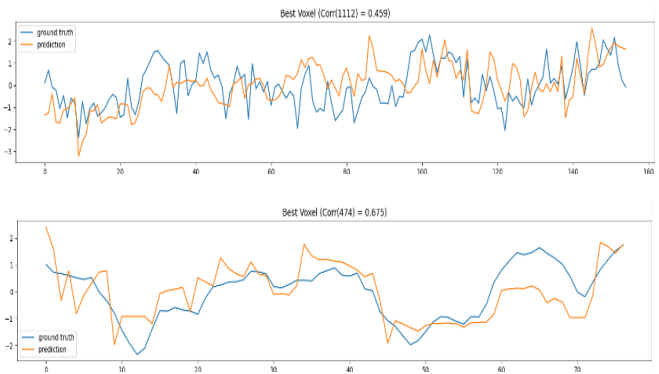
Improvement of the Encoder

- **1 subject → average activity across 30 subjects**
- batch size = 1 → batch size = 16
- other hyperparameters

Encoder 2.0

Global



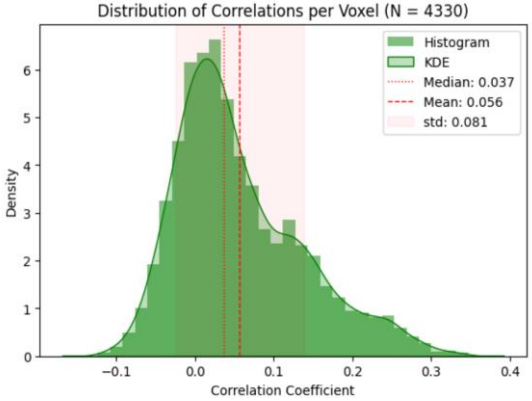


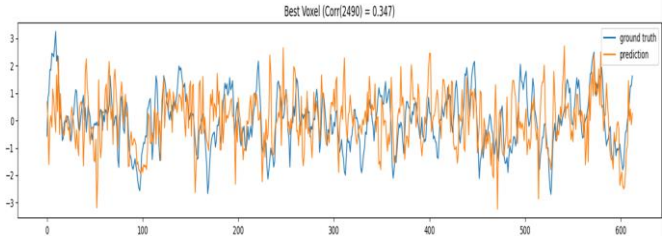
Correlations over time

How well each voxel has been predicted

→ Mean correlations: **0.061**
→ Median correlations: 0.060

You Again



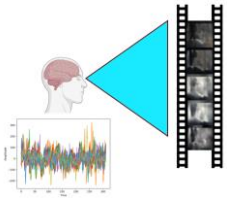


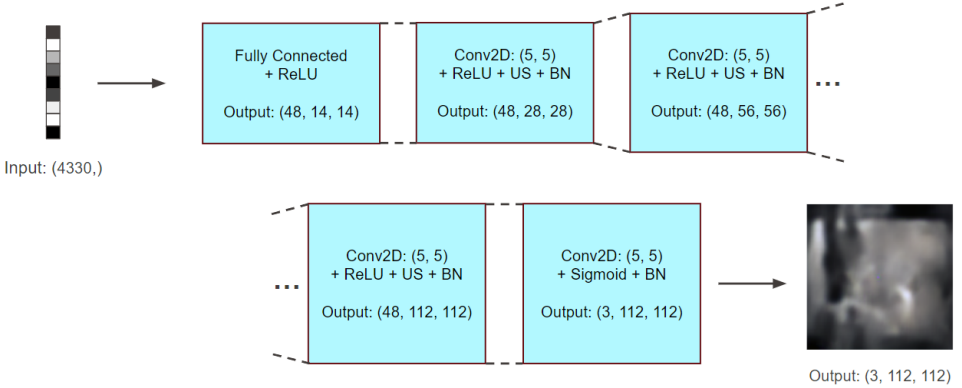
Correlations over time

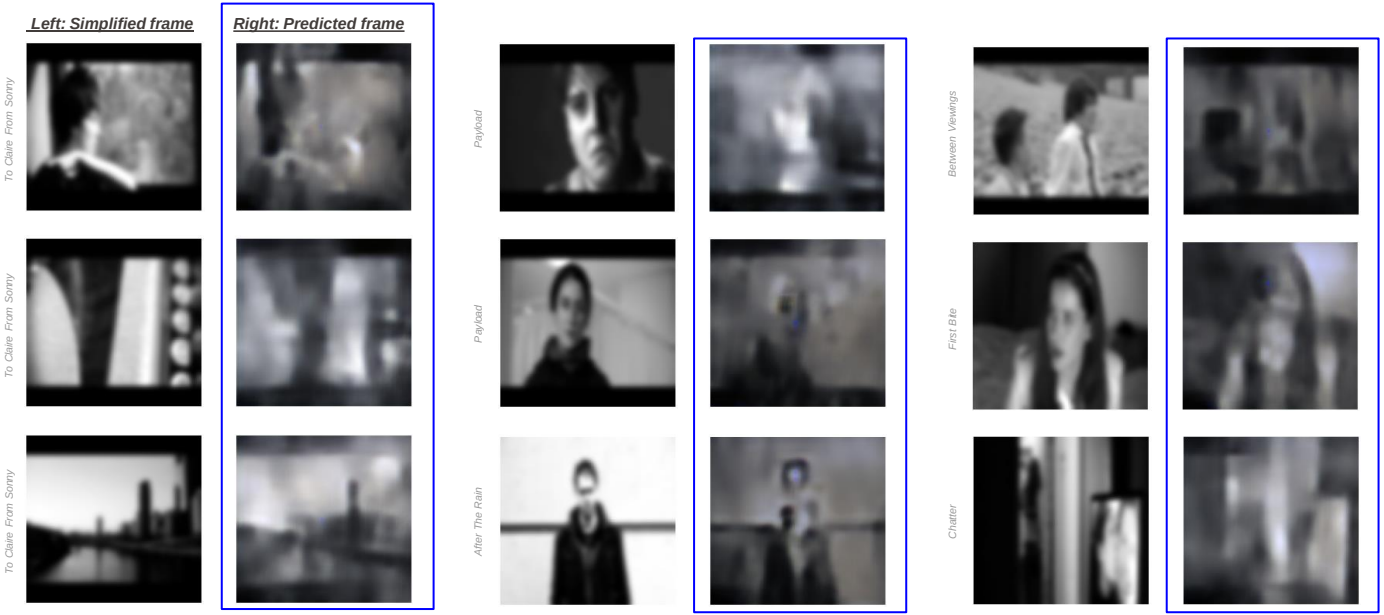
How well each voxel has been predicted

→ Mean correlations: **0.056**
→ Median correlations: 0.037

Decoder

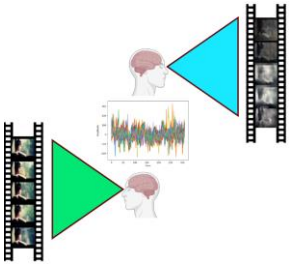






→ The decoder can be seen as a regularizer for the encoder

Encoder-Decoder

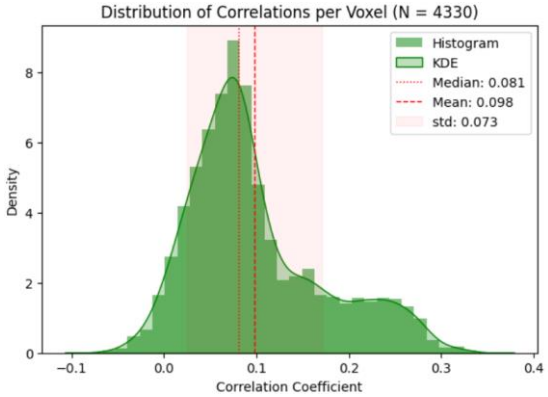


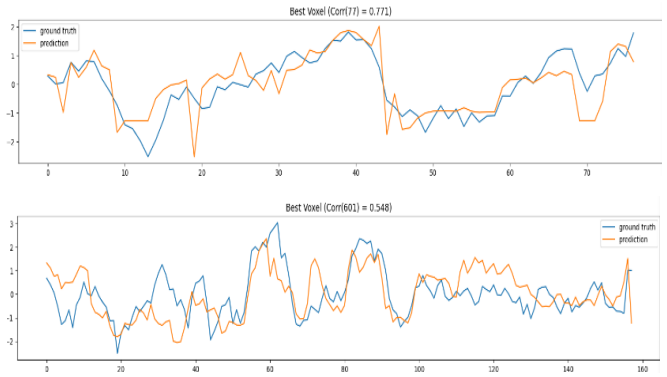
Model training

- End-to-end training: all the weights are updated at each iteration
- Objective:
  - minimize the encoder loss
    - minimize **MSE**
    - minimize **cosine distance**
  - minimize the decoder loss
    - maximize **perceptual similarity**
    - maximize **structural similarity**
    - minimize **total variation**

Encoder-Decoder

Global



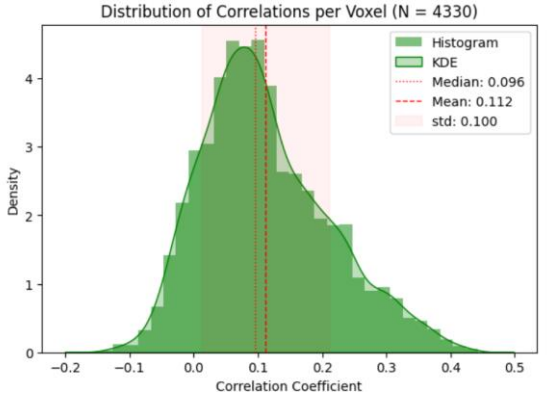


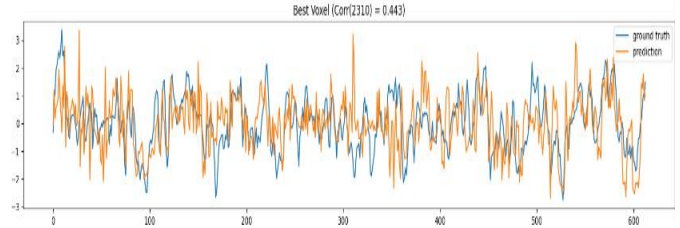
Correlations over time

How well each voxel has been predicted

→ Mean correlations: **0.098**
→ Median correlations: 0.081

You Again





Correlations over time

How well each voxel has been predicted

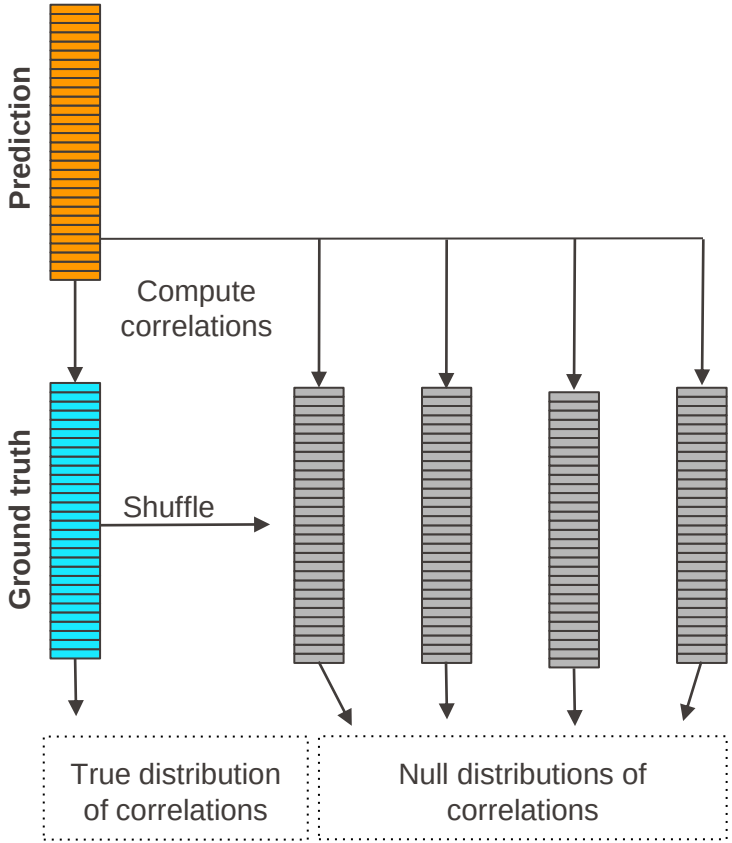
→ Mean correlations: **0.112**
→ Median correlations: 0.96

Encoder-Decoder

Global

You Again

Significance – Permutation test 101



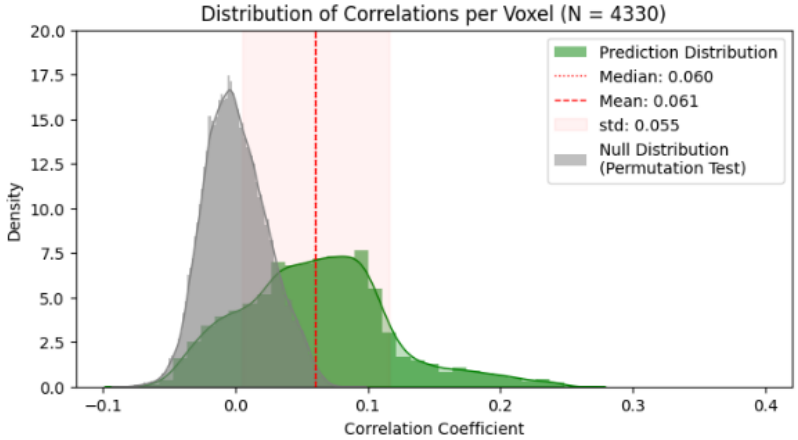
For each voxel:

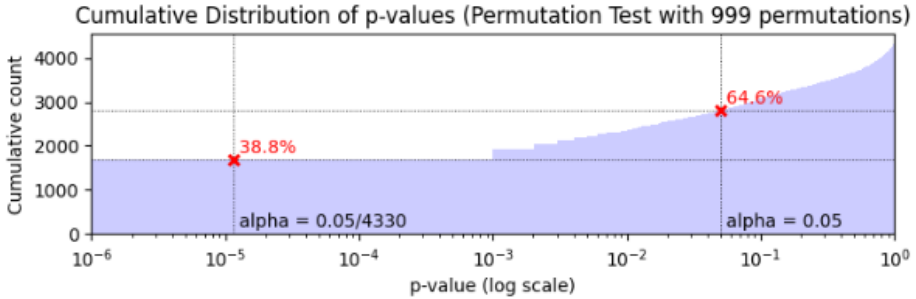
- compute correlations between prediction and ground truth
- compute correlations between prediction and shuffled ground truth multiple times
- p-value: calculate the probability that the true correlation value is lower than the correlation with a shuffled label

→ **get 1 p-value per voxel**

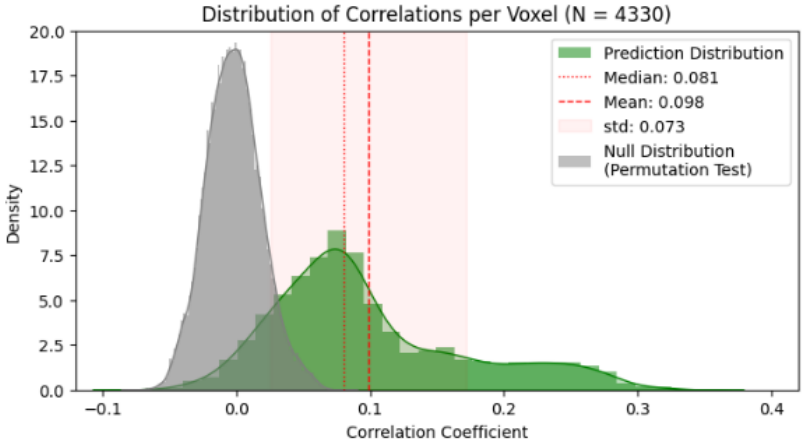
Significance – Global

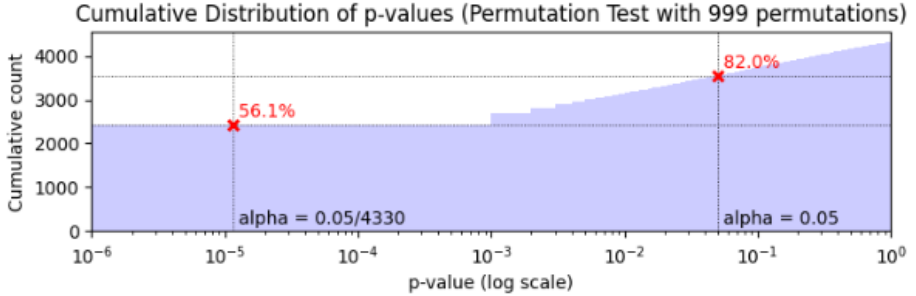
Encoder





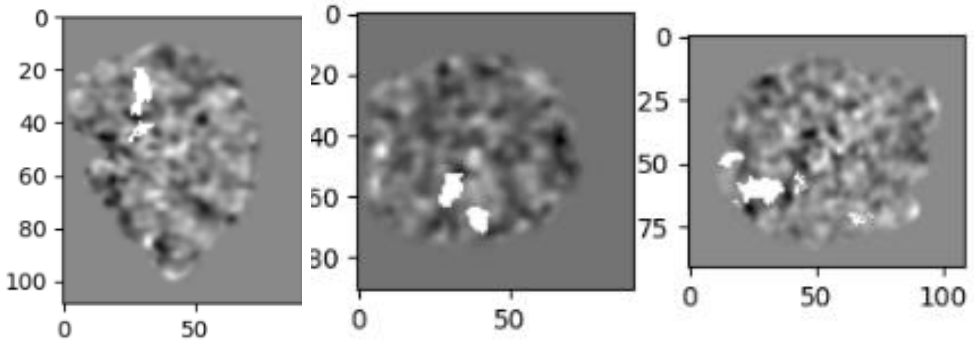
Encoder-Decoder

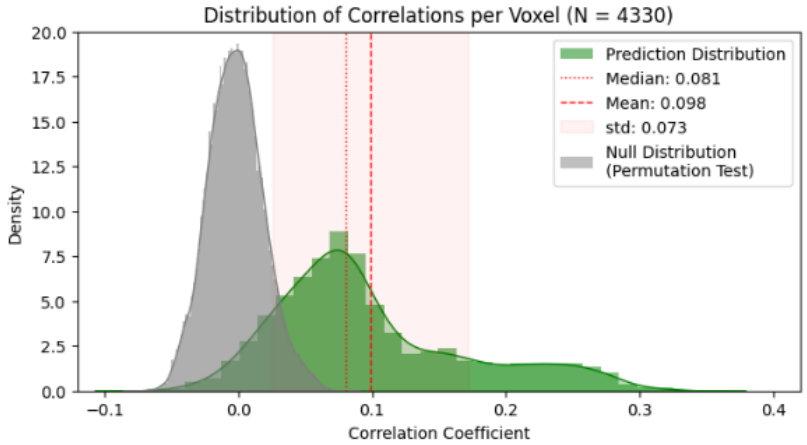




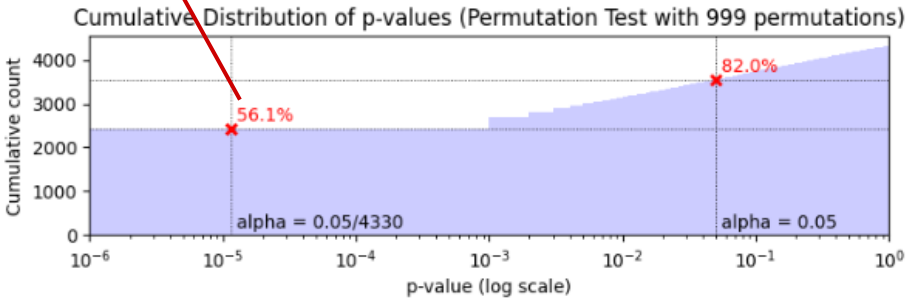
[DETOUR] Significance – Localisation

Encoder-Decoder



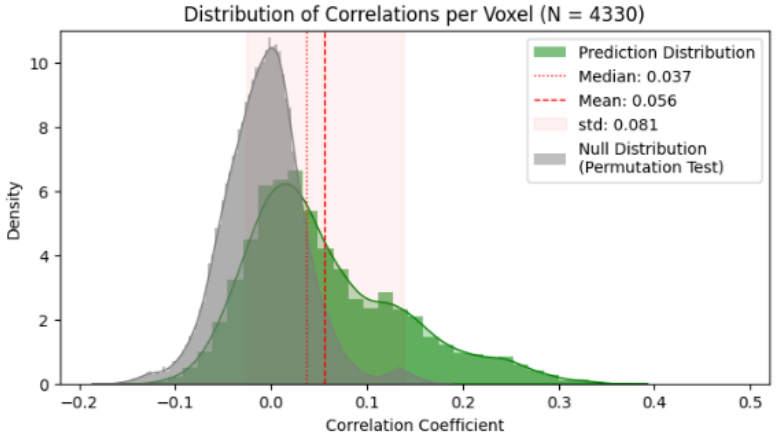


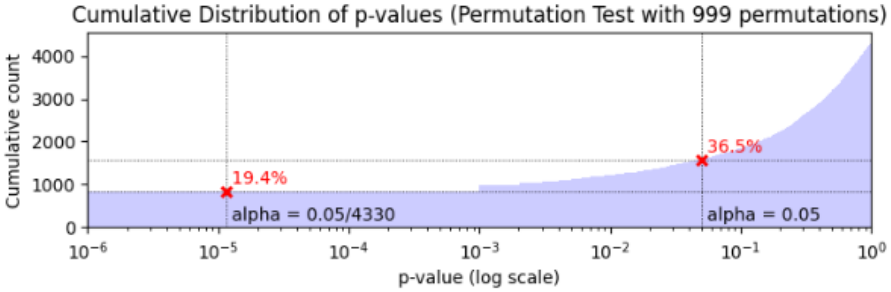
Where are located the most significant voxels?



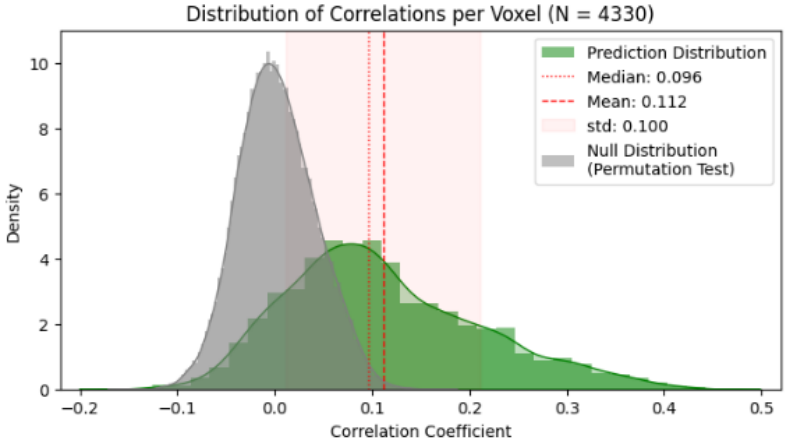
Significance – You Again

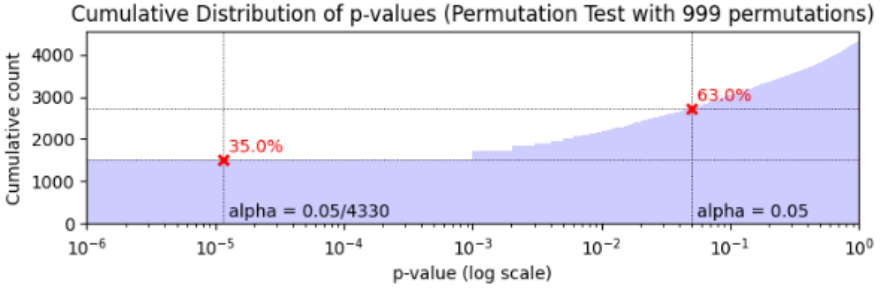
Encoder



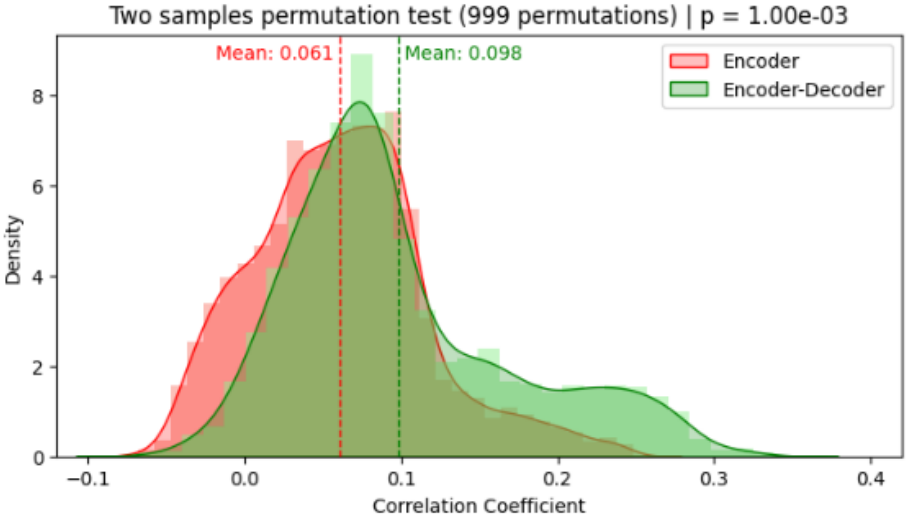


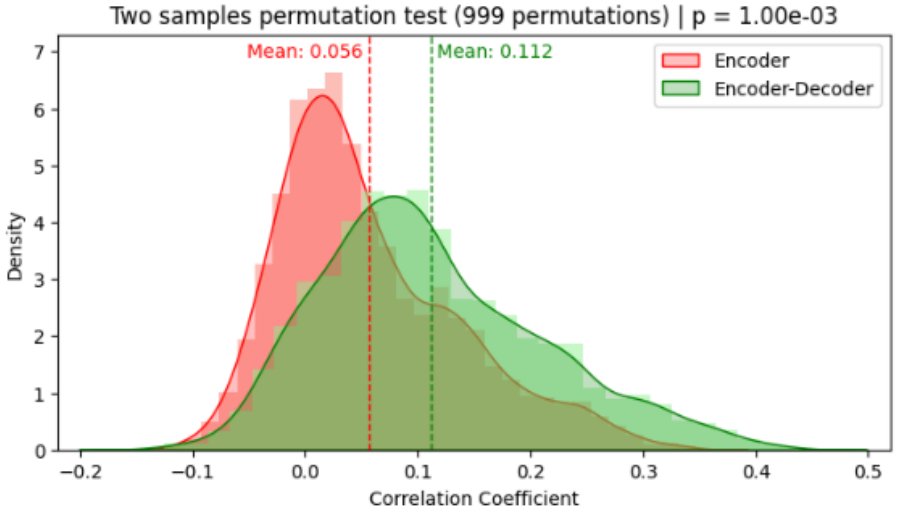
Encoder-Decoder





Significance – Model comparison





Global

→ Means difference: 0.037
→ p < 0.05

You Again

→ Means difference: 0.056
→ p < 0.05

Scientific aim(s)

ML4Science results
→ Average correlations on the **test set**: 0.001

Can we improve it?

1. Build an **encoder** to predict brain activity while someone is watching a movie

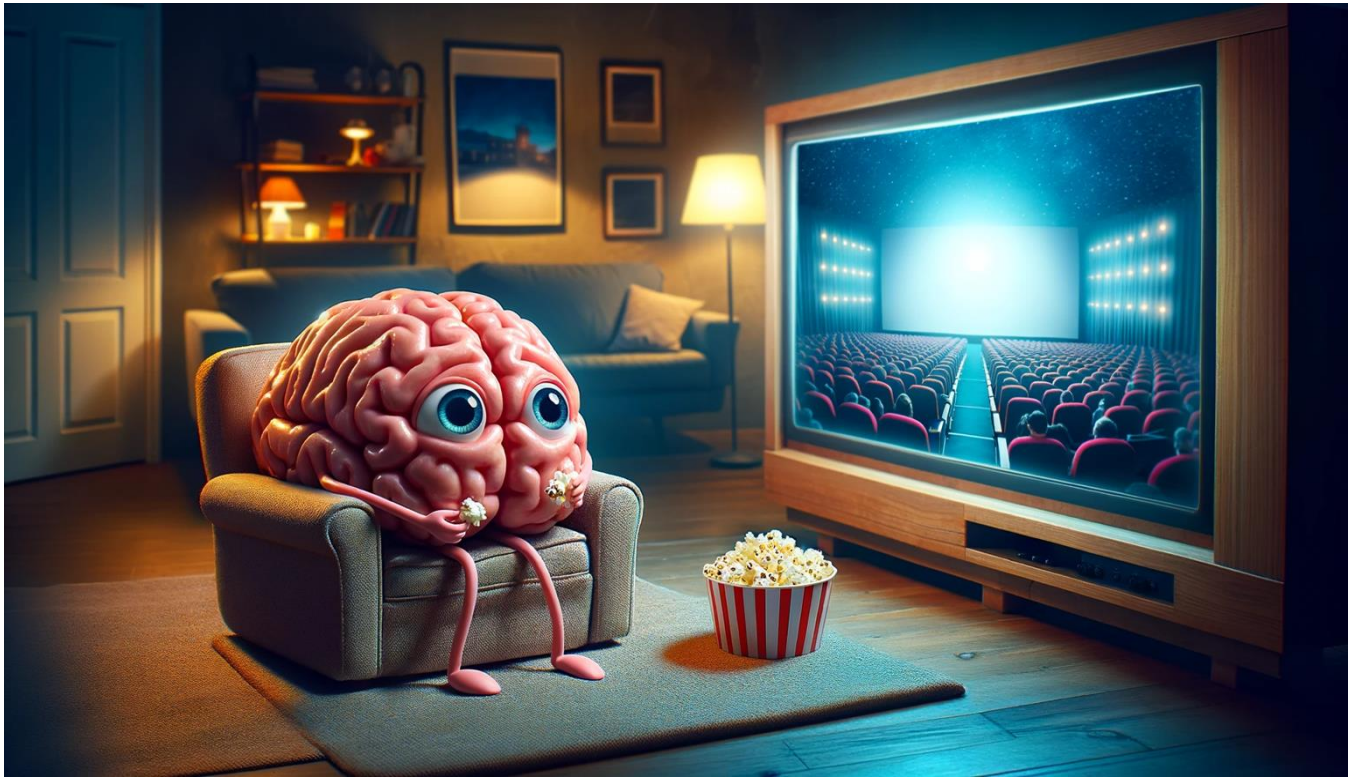
→ The encoder alone can predict brain activity on an **unseen movie** with an average correlation of **0.056**

2. Build an **decoder** to validate the predicted brain activity

→ The decoder alone can learn to reconstruct movie frames from brain activity

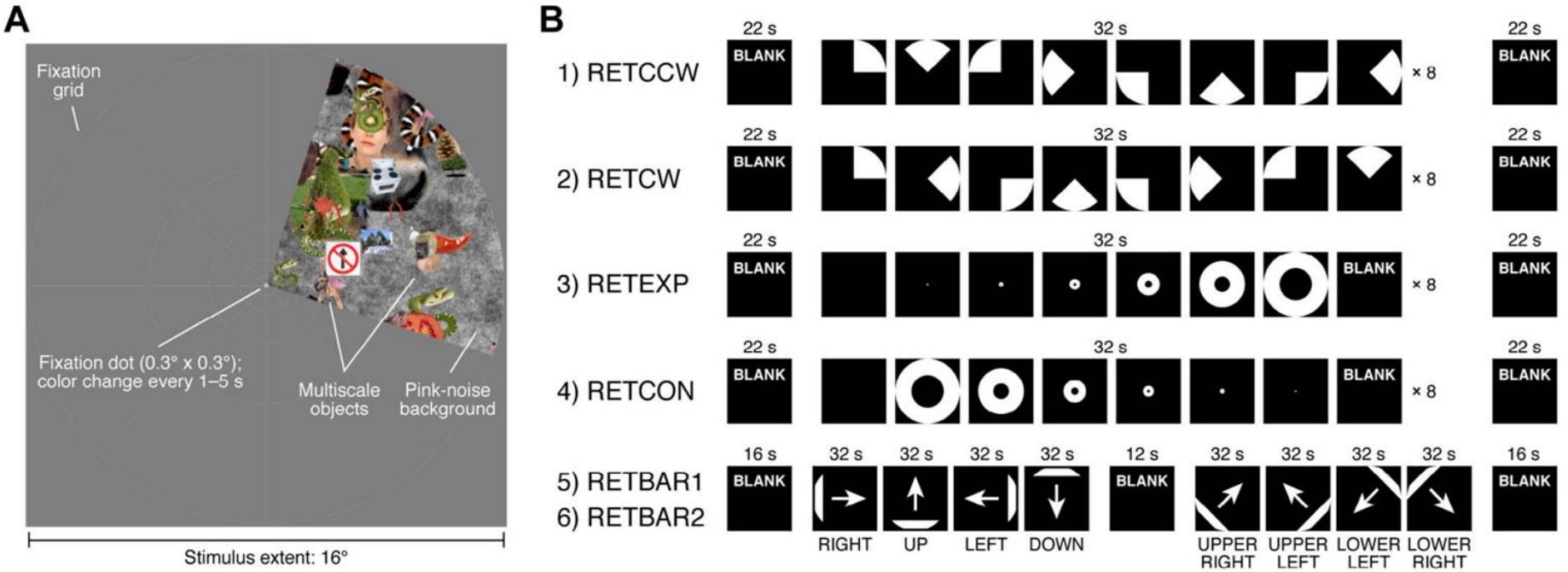
3. Merge both models into an end-to-end **encoder-decoder** to generate fMRI data on unseen movies

→ The end-to-end encoder-decoder can predict brain activity on an **unseen movie** with an average correlation of **0.112**

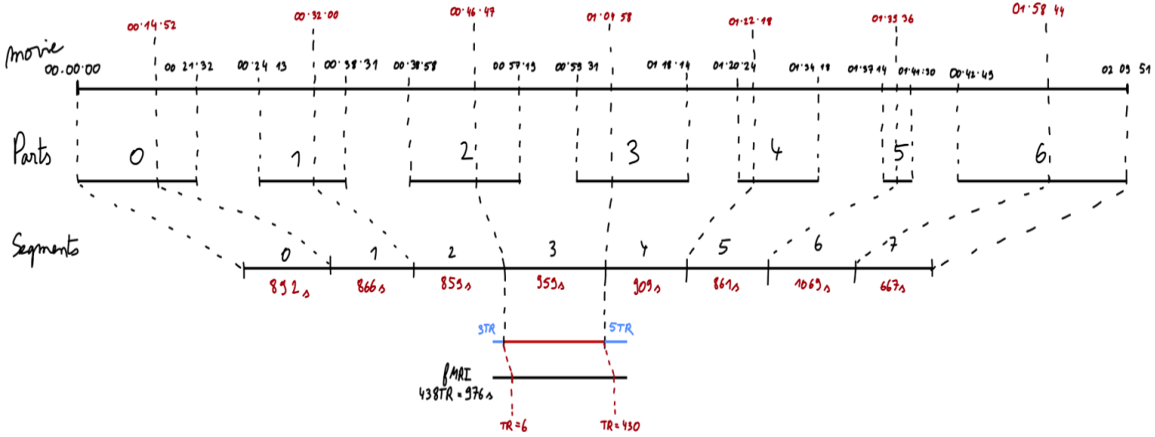


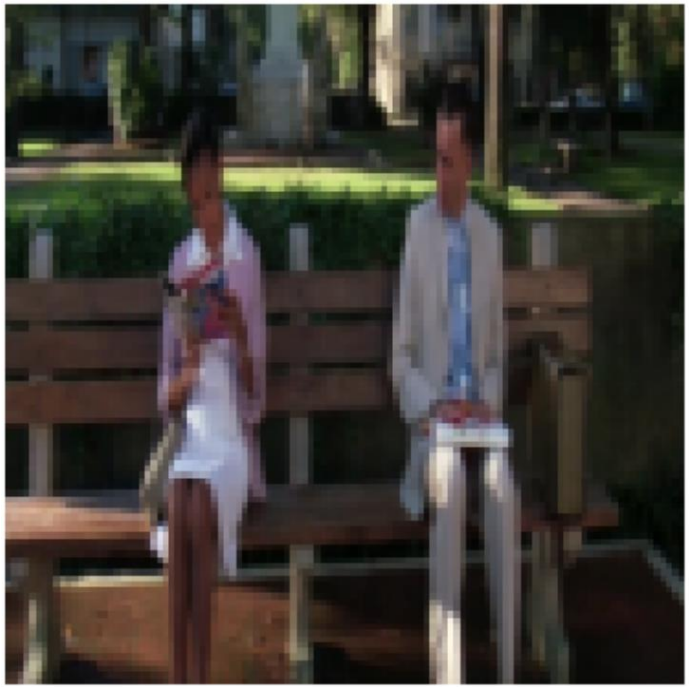
Thank you for your attention!

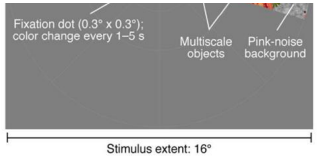
Bonus: what went wrong…

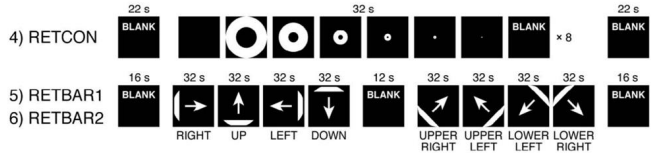


Bonus: what went wrong...

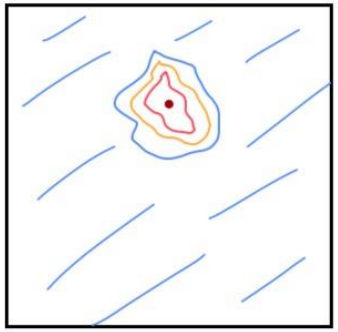


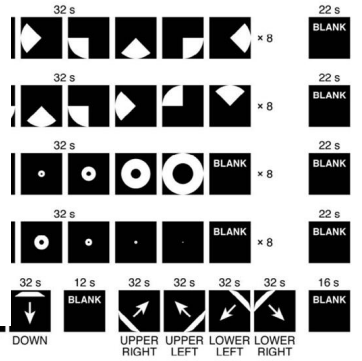


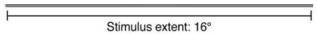




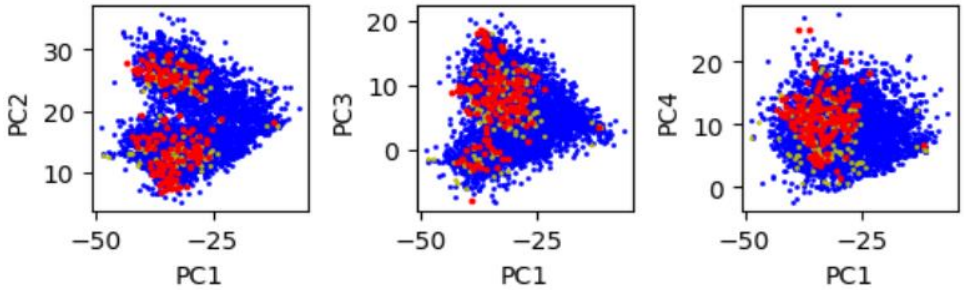
Bonus: what went wrong...

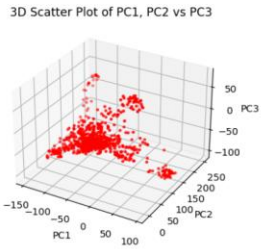


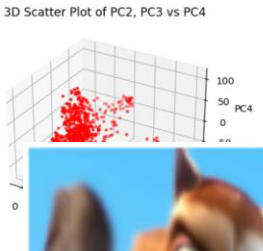


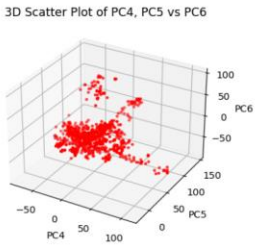


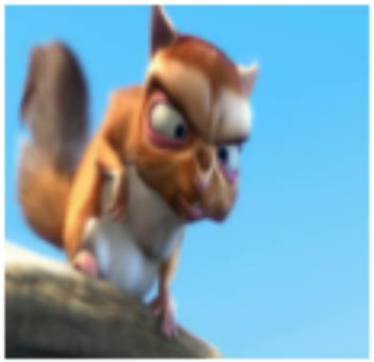
Bonus: what went wrong...

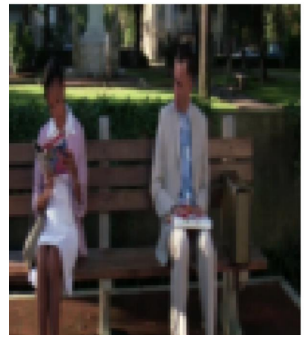


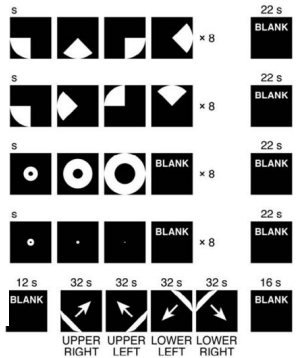








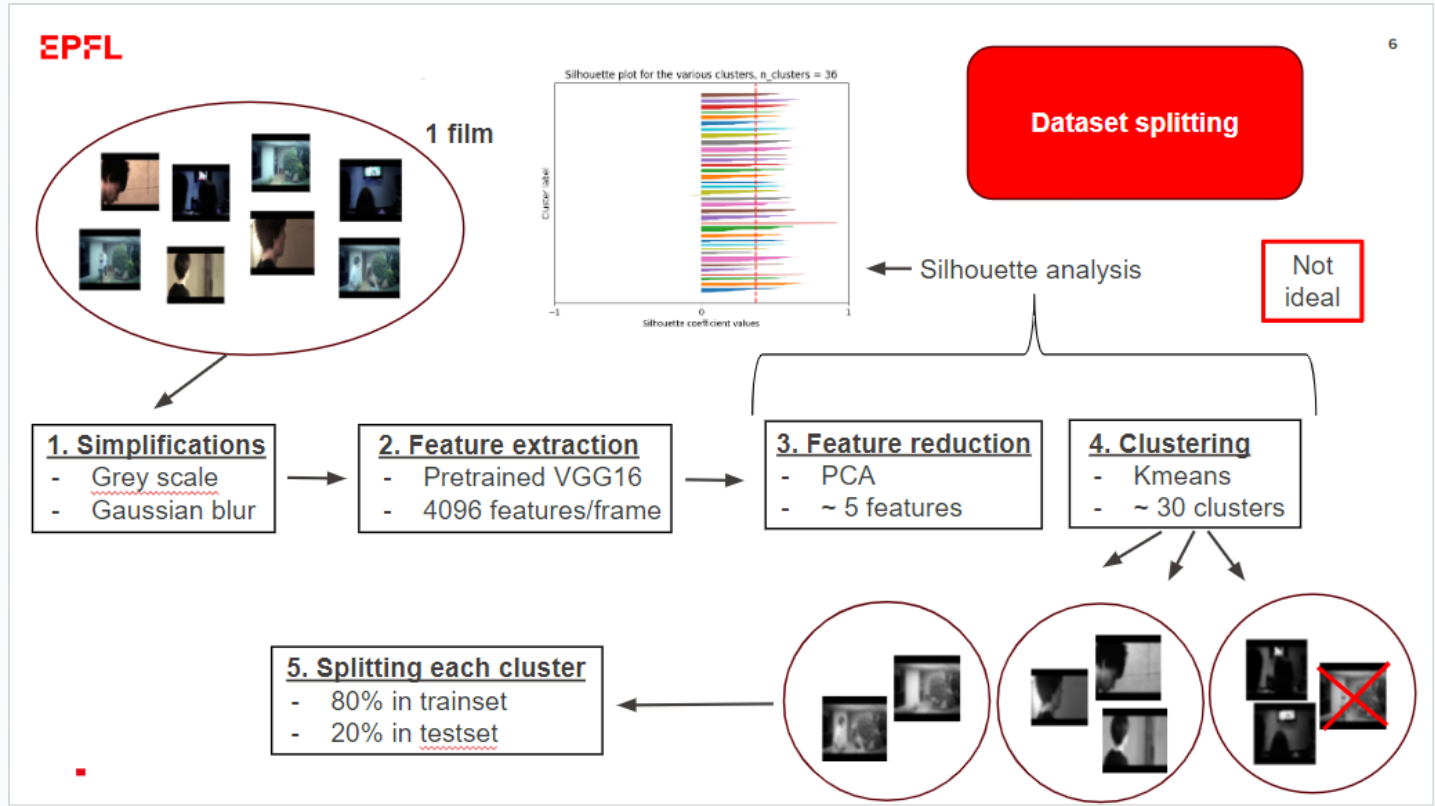


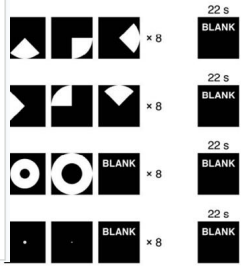


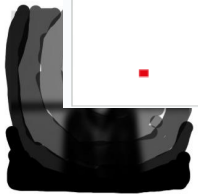
Stimulus extent: 16°

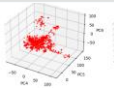
DOWN

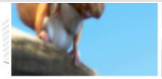
Bonus: what went wrong...

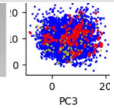


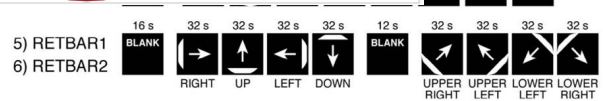




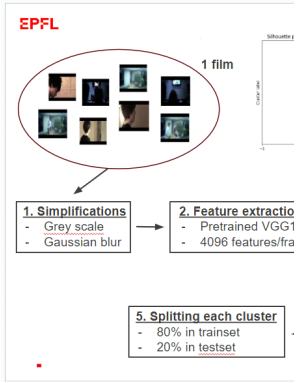


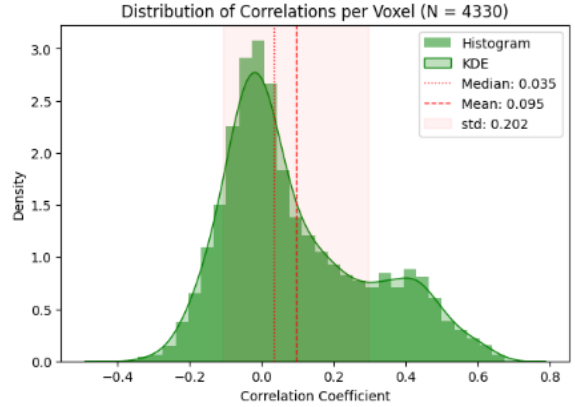


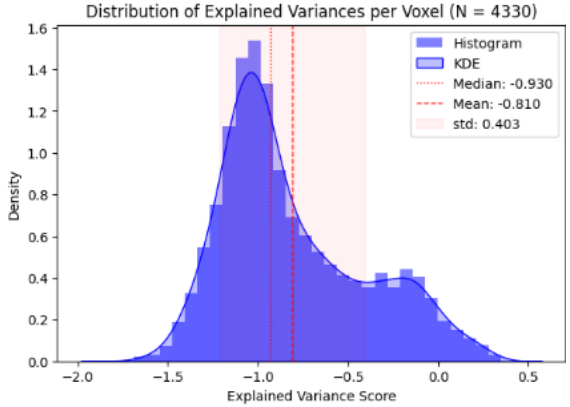


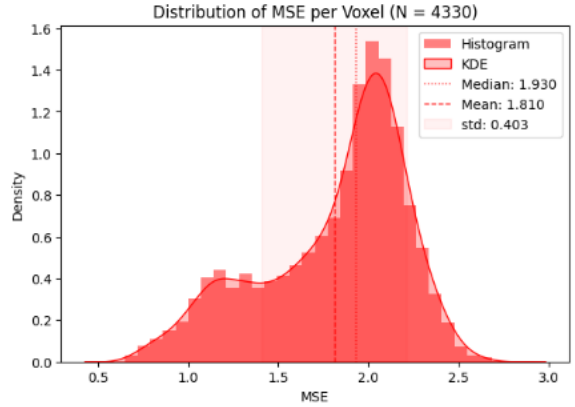


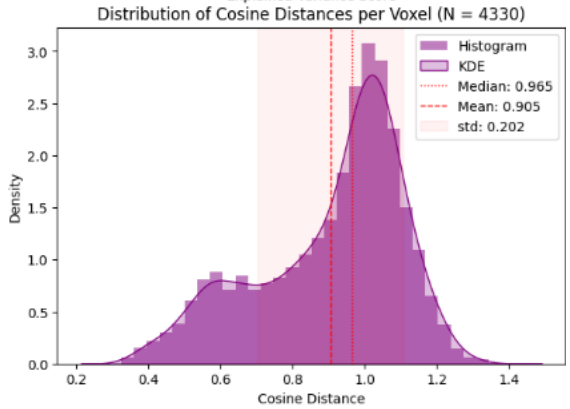
Bonus: what went wrong...

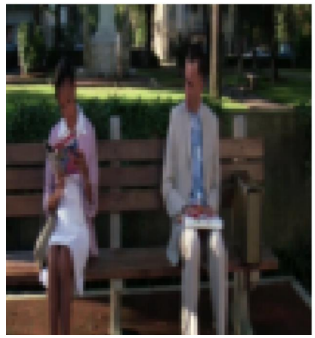


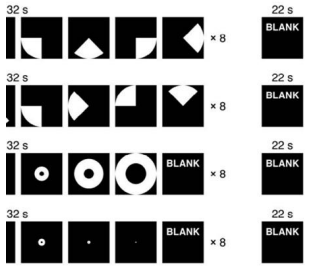


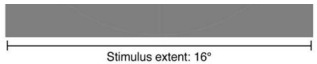




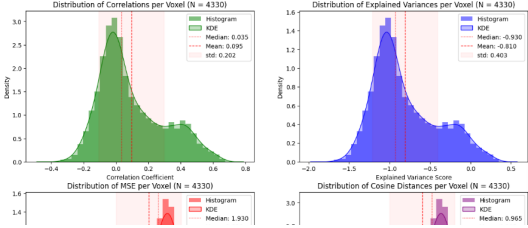


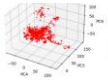


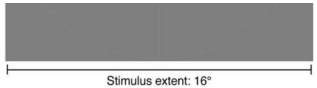


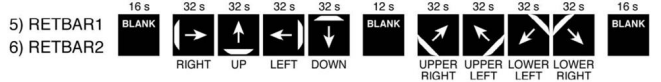


Bonus: what went wrong...

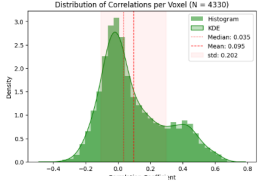


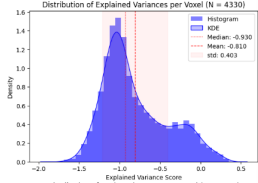


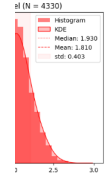


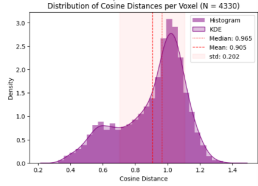


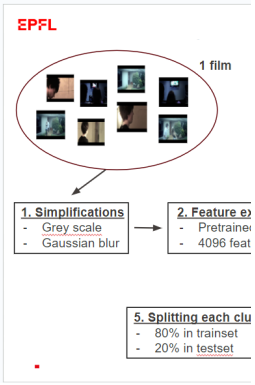
went wrong...

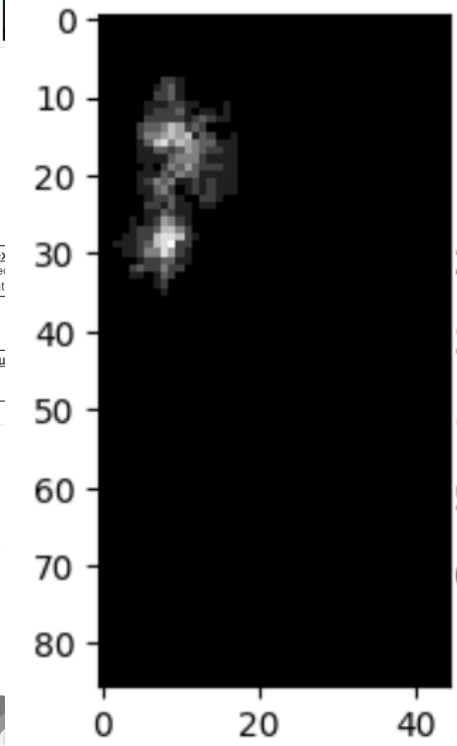

Distribution of Correlations per Voxel (N = 4330)
Histogram
KDE
Median: 0.035
Mean: 0.095
std: 0.202
Density

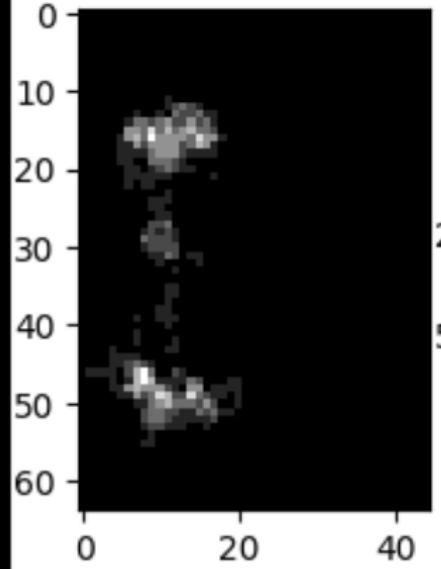

Distribution of Explained Variances per Voxel (N = 4330)
Histogram
KDE
Median: -0.930
Mean: -0.810
std: 0.403
Density
Explained Variance Score

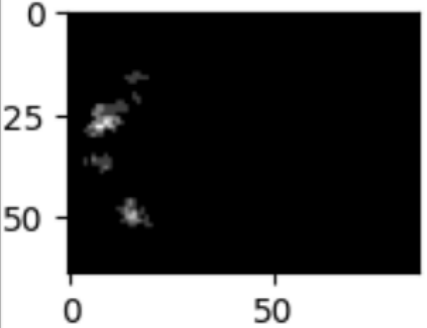

Histogram
KDE
Median: 1.930
Mean: 1.810
std: 0.403

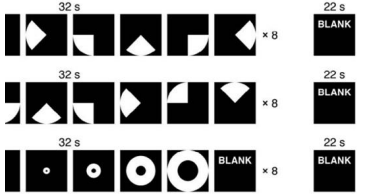

Distribution of Cosine Distances per Voxel (N = 4330)
Histogram
KDE
Median: 0.965
Mean: 0.905
std: 0.202
Density
Cosine Distance

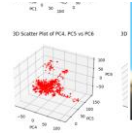

EPFL
1 film
1. Simplifications
- Grey scale
- Gaussian blur
2. Feature ex
- Pretrainec
- 4096 feat
5. Splitting each clu
- 80% in trainset
- 20% in testset

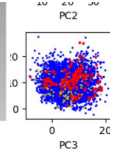


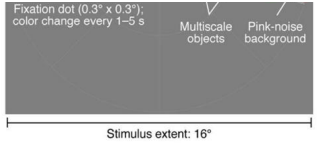


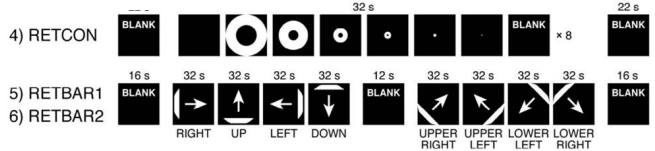



32 s
22 s
BLANK
× 8
32 s
22 s
BLANK
× 8
32 s
22 s
BLANK
BLANK
× 8

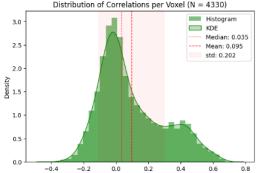


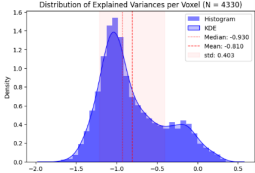

PC2
PC3

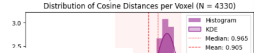

Fixation dot (0.3° x 0.3°); color change every 1–5 s
Multiscale objects
Pink-noise background
Stimulus extent: 16°

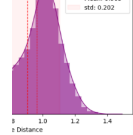

4) RETCON
5) RETBAR1
6) RETBAR2
BLANK
32 s
22 s
16 s
12 s
× 8
RIGHT
UP
LEFT
DOWN
UPPER RIGHT
UPPER LEFT
LOWER LEFT
LOWER RIGHT

went wrong...



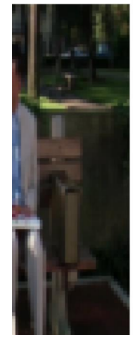




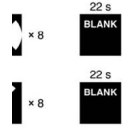


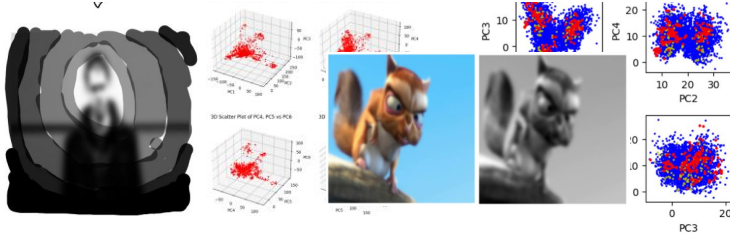
Delete Failed

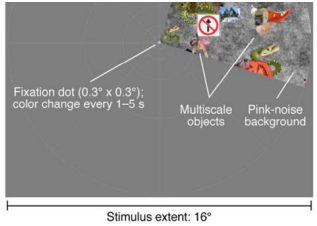
send2trash f`1ailed: [Errno 28] No space left on device

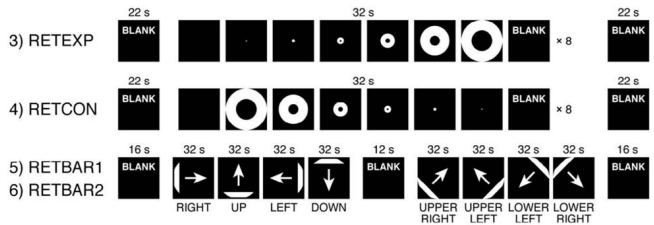


Dismiss





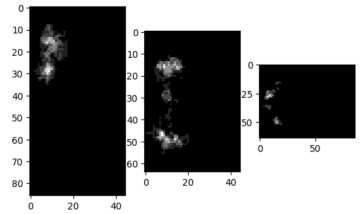


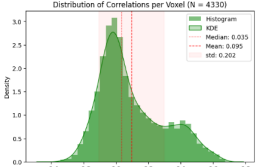


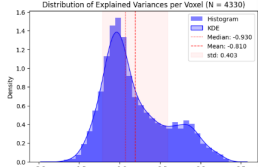
Delete Failed

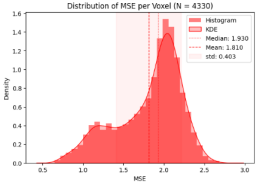
send2trash f`1ailed: [Errno 28] No space left on device

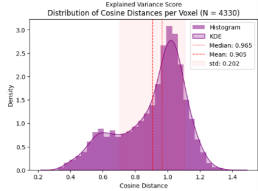
Dismiss

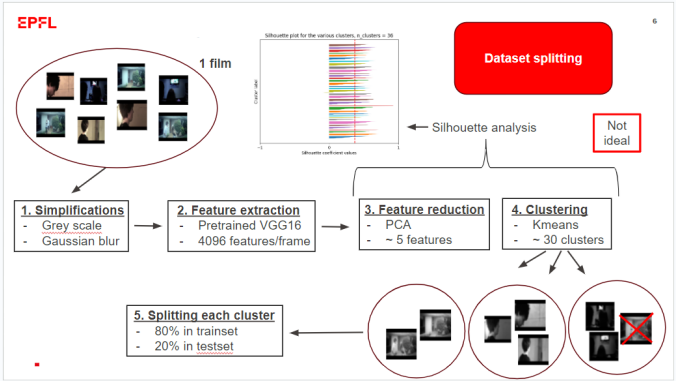


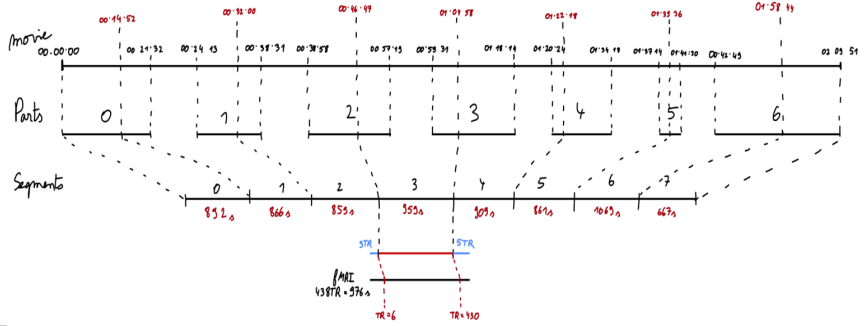


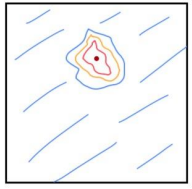


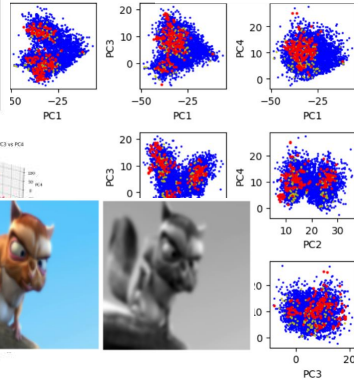


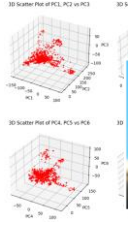


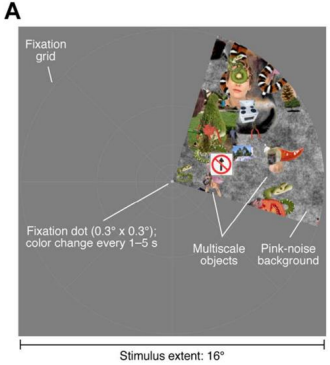


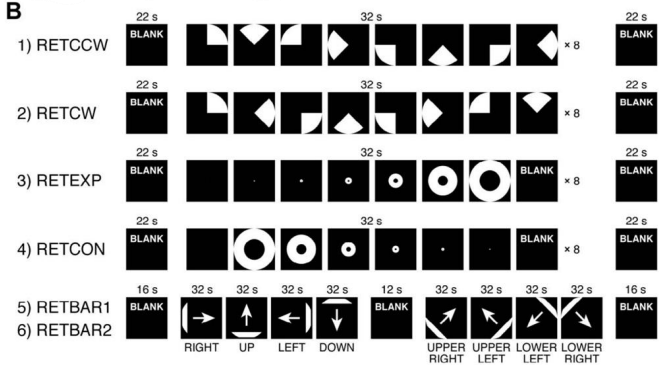


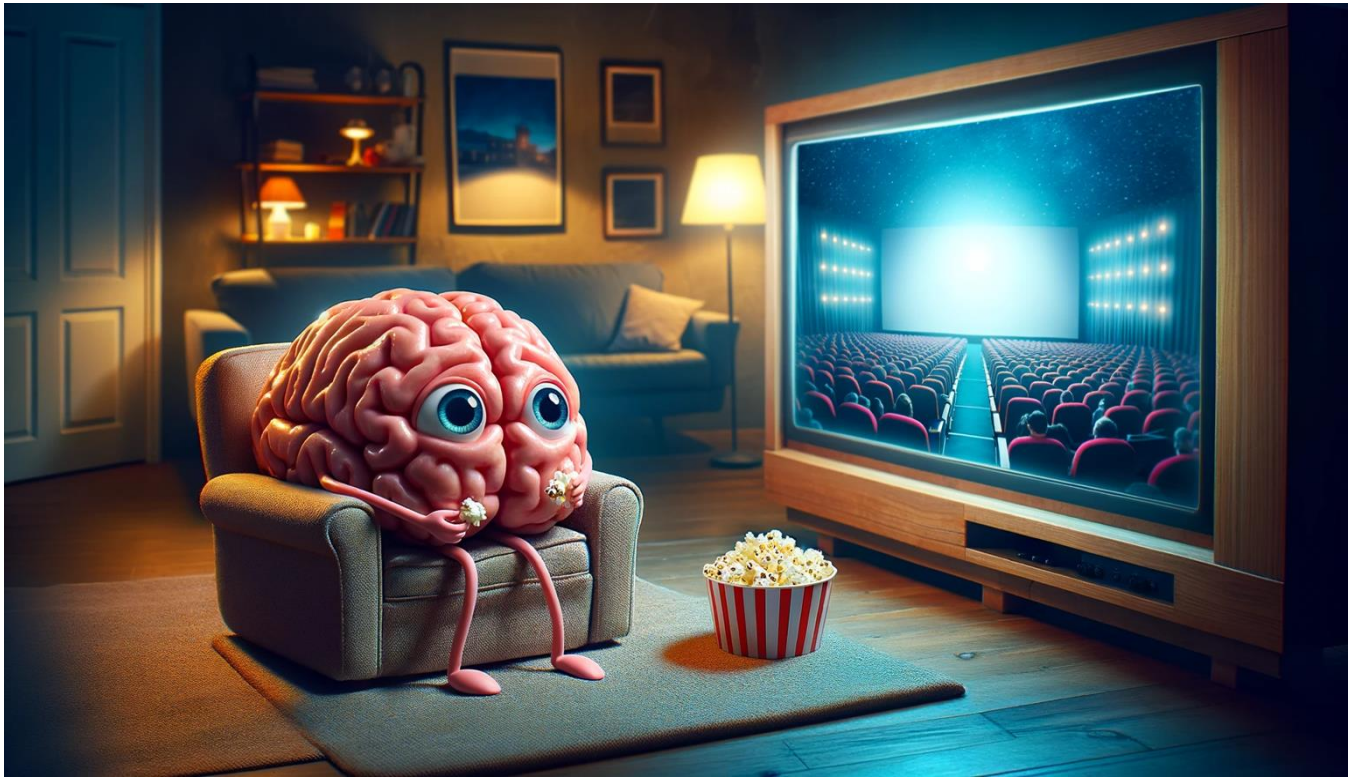












Thank you for your attention!