Brain decoding with Machine Learning

Leila Reddy
CerCo/CNRS

Outline

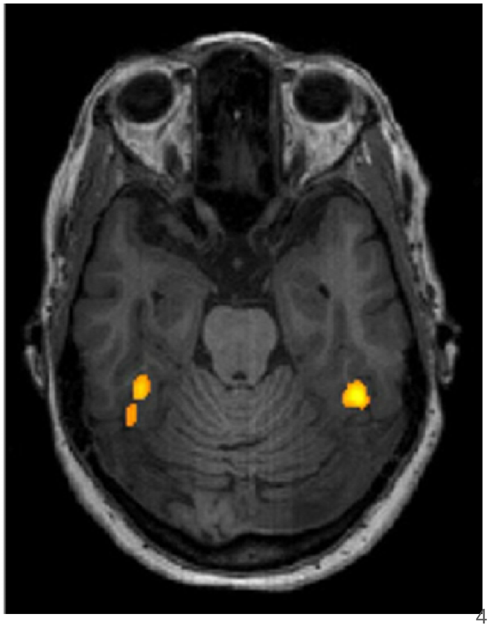
- Brief introduction to fMRI
- The “old” way: decoding stimulus info using classification methods (e.g., SVM).
- Decoding visual and language stimuli using DNN features
- Face and scene reconstruction using DNN latent spaces

Brain recording methods

- Single neuron electrophysiology:
  - invasive, mostly in animals, rare in humans
- EEG (Electroencephalography)
  - Measure electrical activity from the scalp
    - Good temporal resolution, poor spatial resolution
- MEG (Magnetoencephalography)
  - Similar to EEG but records magnetic fields
- fMRI (functional magnetic resonance imaging)

Brain Decoding : fMRI

- Functional magnetic resonance imaging (fMRI):
  - powerful tool to study the human brain non-invasively.
  - measures brain activity by detecting changes in blood flow.
  - resolution:
    - not neurons but voxels (like 3D pixels)
    - ~10^5 neurons in one $3mm^3$ voxel
    - millimeter scale resolution



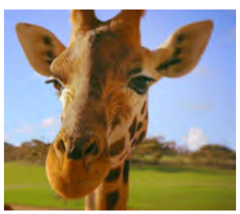
Interpretation of mental contents

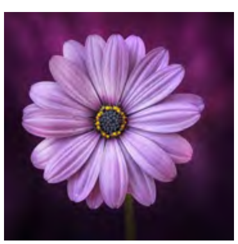
- Brain decoding: decode what people see, imagine, remember etc.
- A decoder is trained to learn a relationship between fMRI patterns and the “content” to be decoded (e.g., a seen image).

Brain Decoding : fMRI patterns

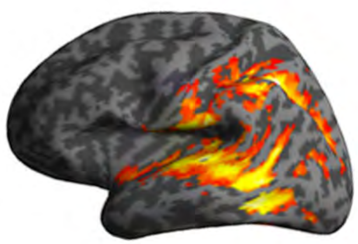
- Input gives rise to activation patterns in the brain
  - The average univariate response may be identical in different conditions.
  - But the fMRI or multi-voxel *patterns* may still carry information.

Image seen by participant in the scanner

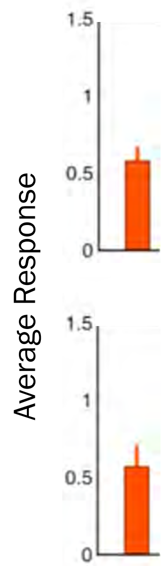




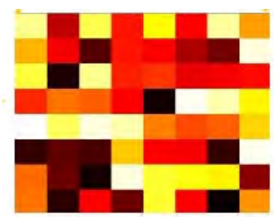
Measure fMRI brain response



fMRI response averaged across voxels

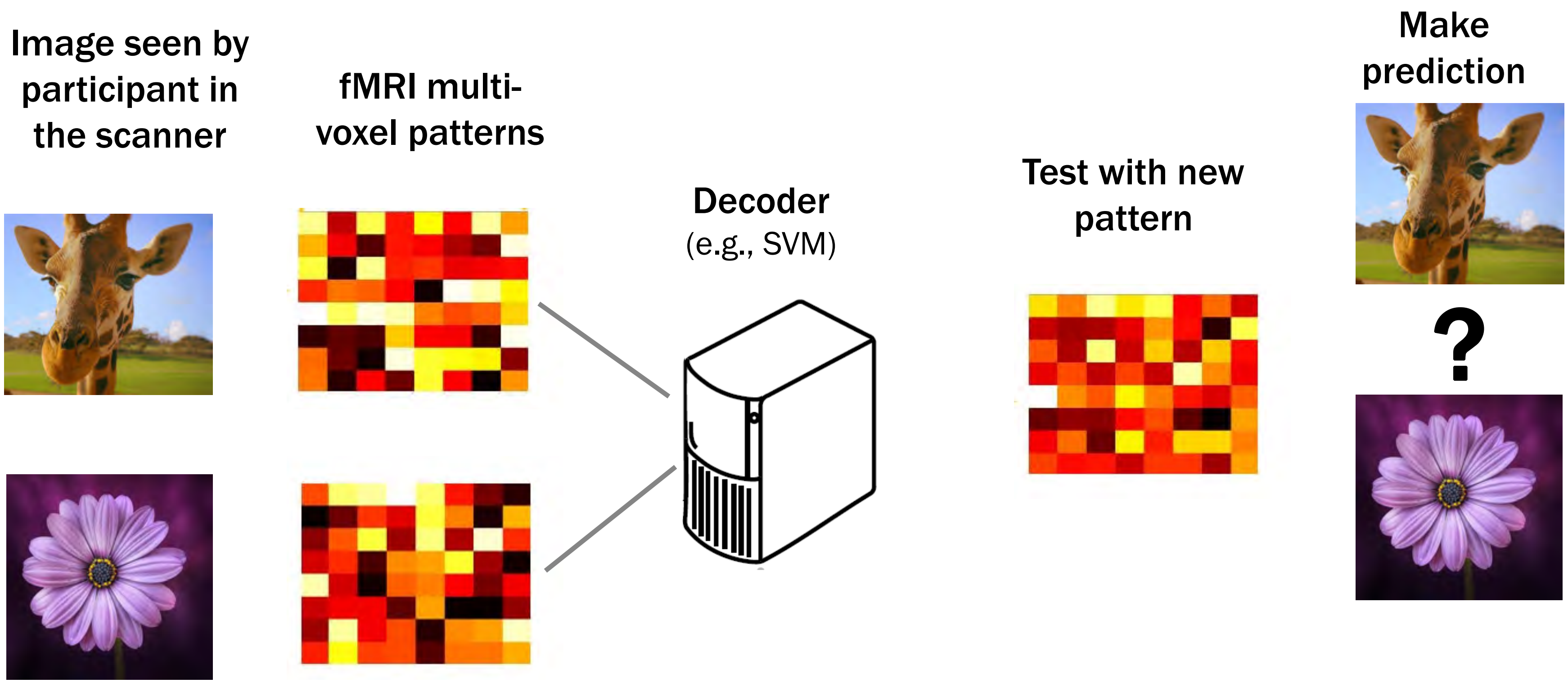


fMRI multi-voxel patterns

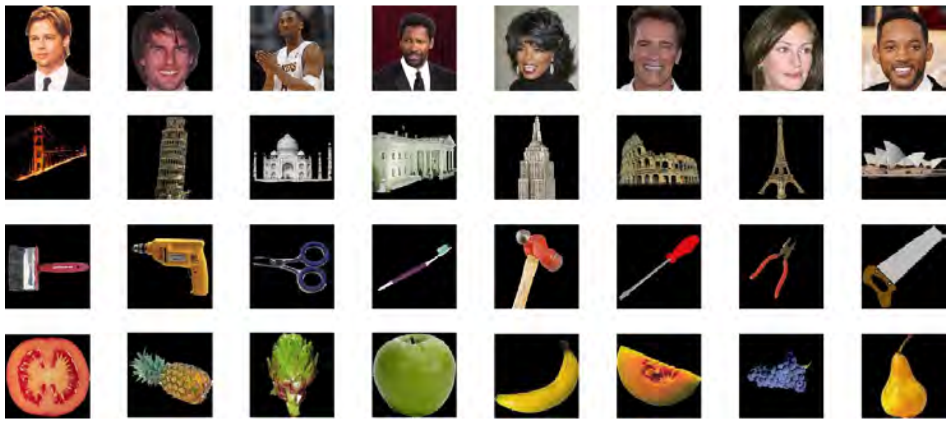


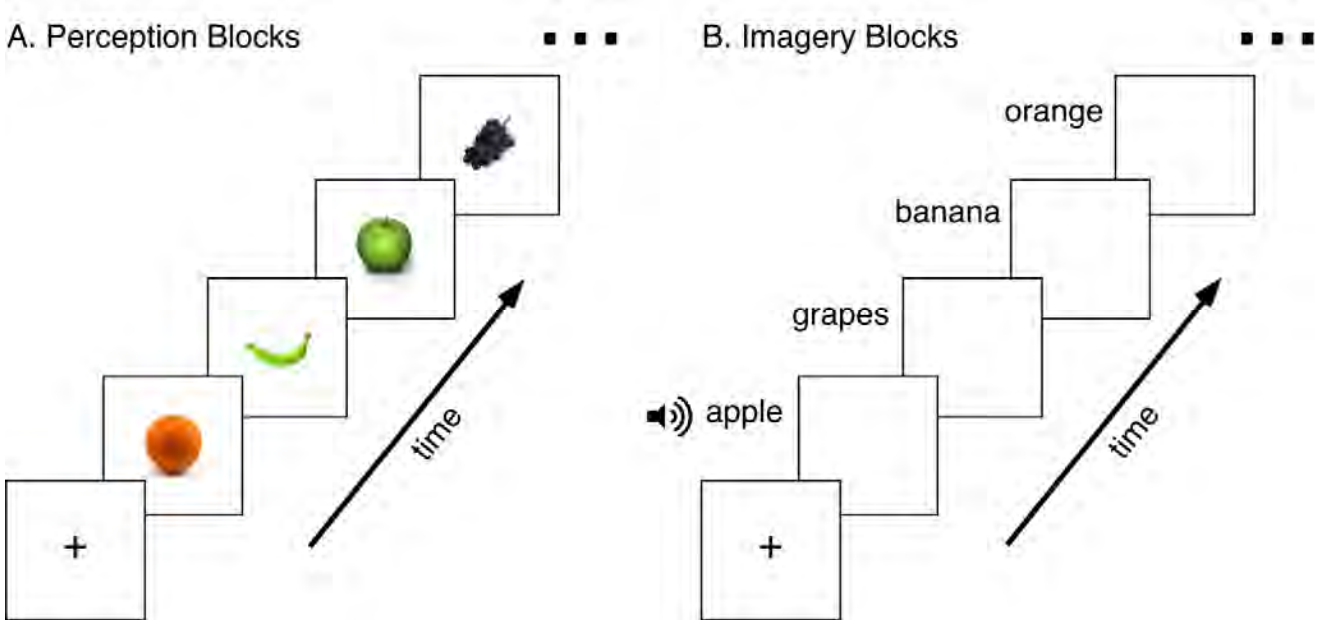
Interpretation of mental contents

- A decoder is trained to learn a relationship between fMRI patterns and the “content” to be decoded (e.g., a seen image).



Decoding perception and imagery

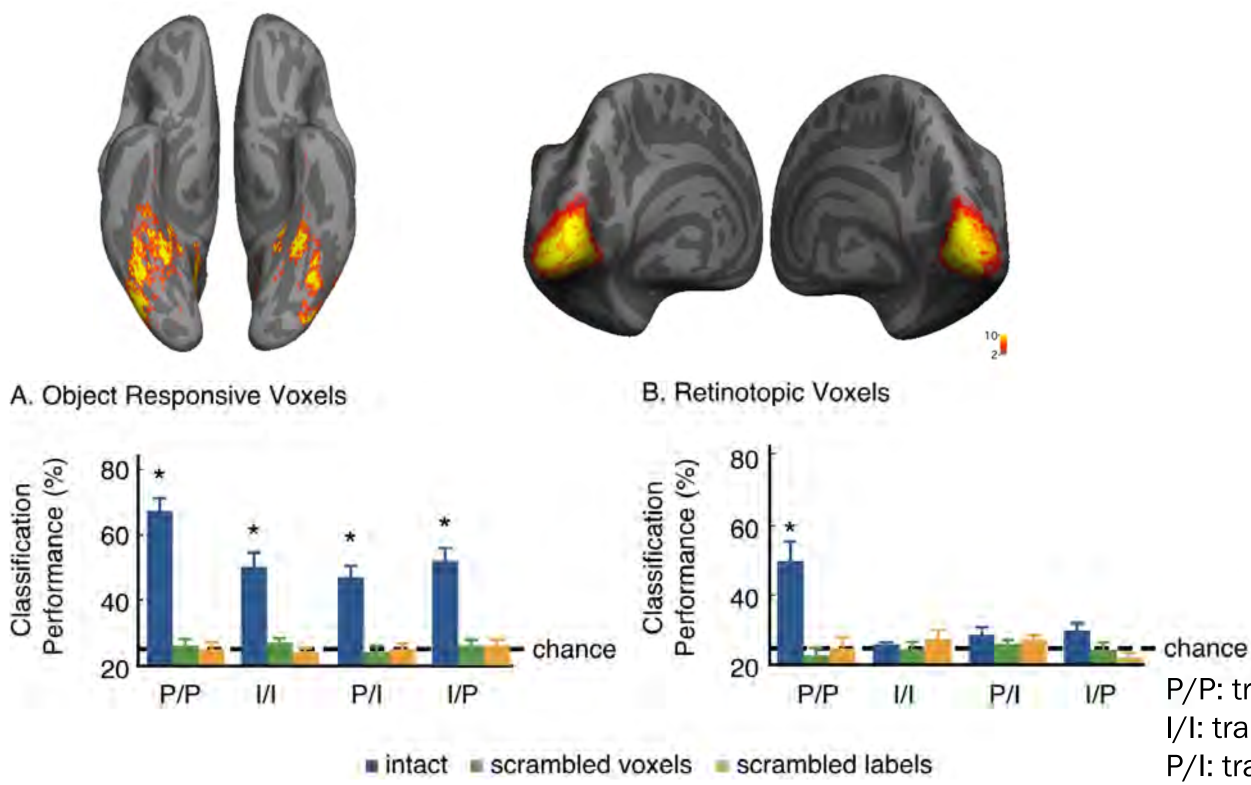




Reddy et al., 2012

Decoding perception and imagery

- Perception decoding from low- and high-level areas.
- Imagery decoding only in high-level regions.
- Perception and imagery share representations in high-level regions.

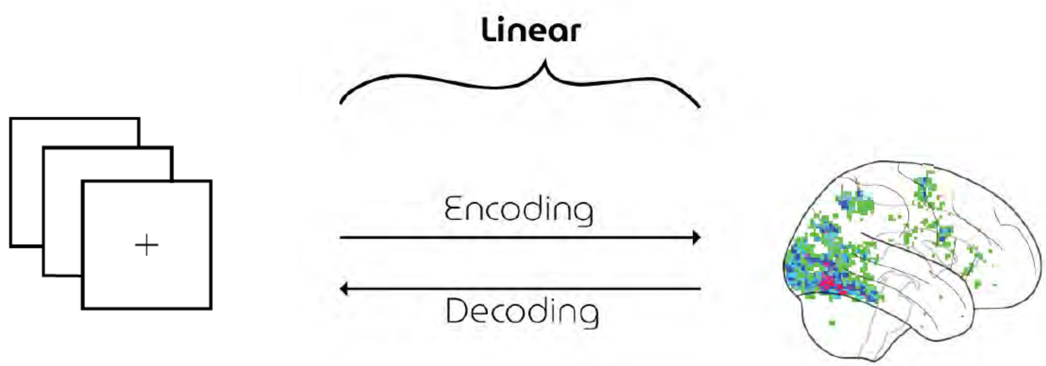


P/P: train and test on perception
I/I: train and test on imagery
P/I: train on perception, test on imagery
I/P: train on imagery, test on perception.

Reddy et al., 2012

DNNs and brain decoding

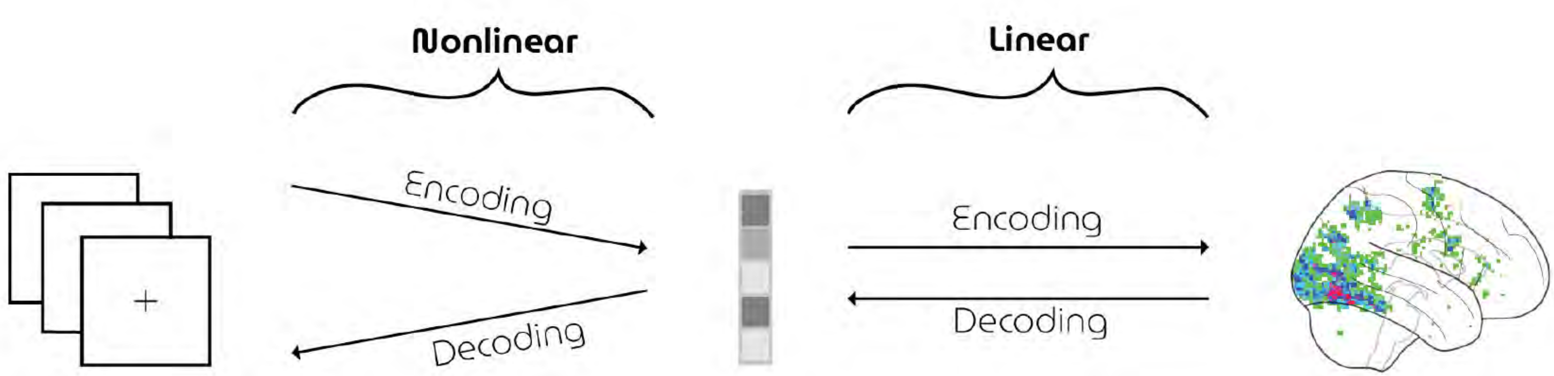
- Classification-based approach (as on previous slide):
  - Decoder learns a mapping between brain patterns and pixel space
  - Limited to decoding the classes used for training
  - Cannot generalize to novel classes



Dado et al., 2020

DNNs and brain decoding

- Classification-based approach (as on previous slide):
  - Decoder learns a mapping between brain patterns and pixel space
  - Limited to decoding the classes used for training
  - Cannot generalize to novel classes



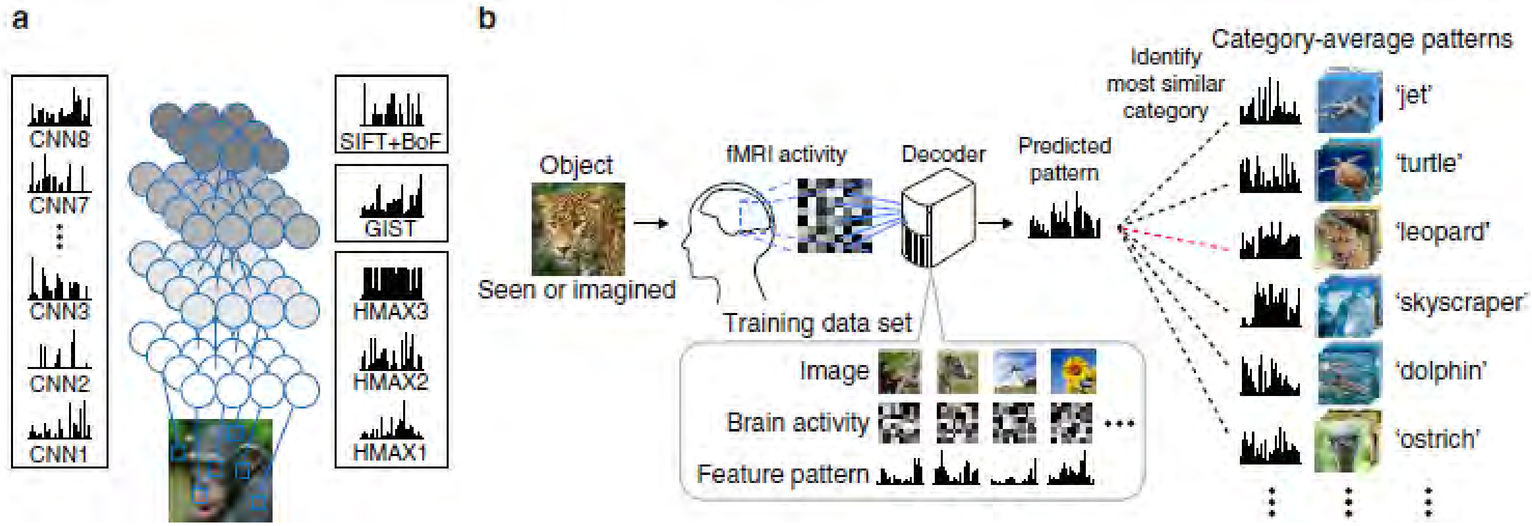
Dado et al., 2020

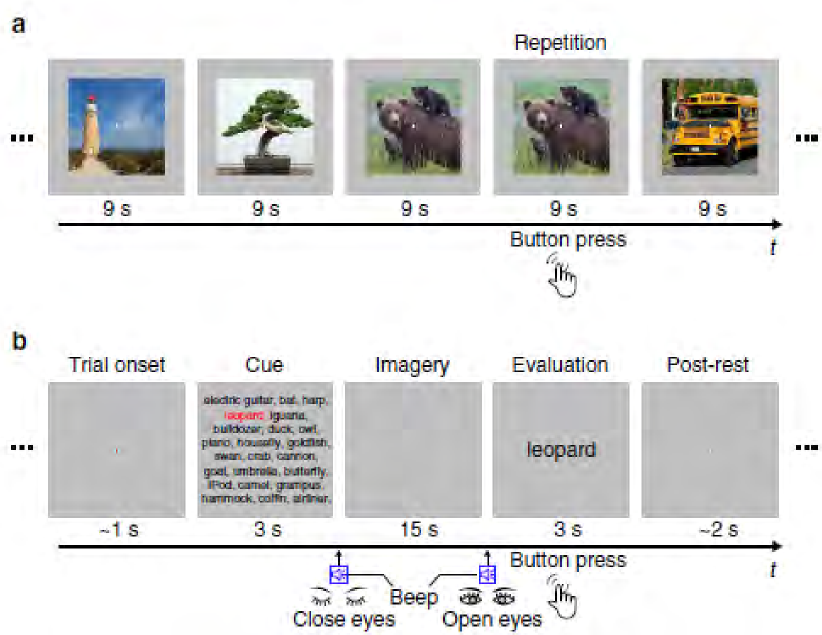
DNNs and brain decoding

- Classification-based approach (as on previous slide):
  - Decoder learns a mapping between brain patterns and pixel space
  - Limited to decoding the classes used for training
  - Cannot generalize to novel classes

- DNNs could provide a model for brain representations
  - Allow generalization to new classes because we can use the feature space in DNN layers
    - Decoder learns a mapping between brain patterns and representations in the feature space of DNNs

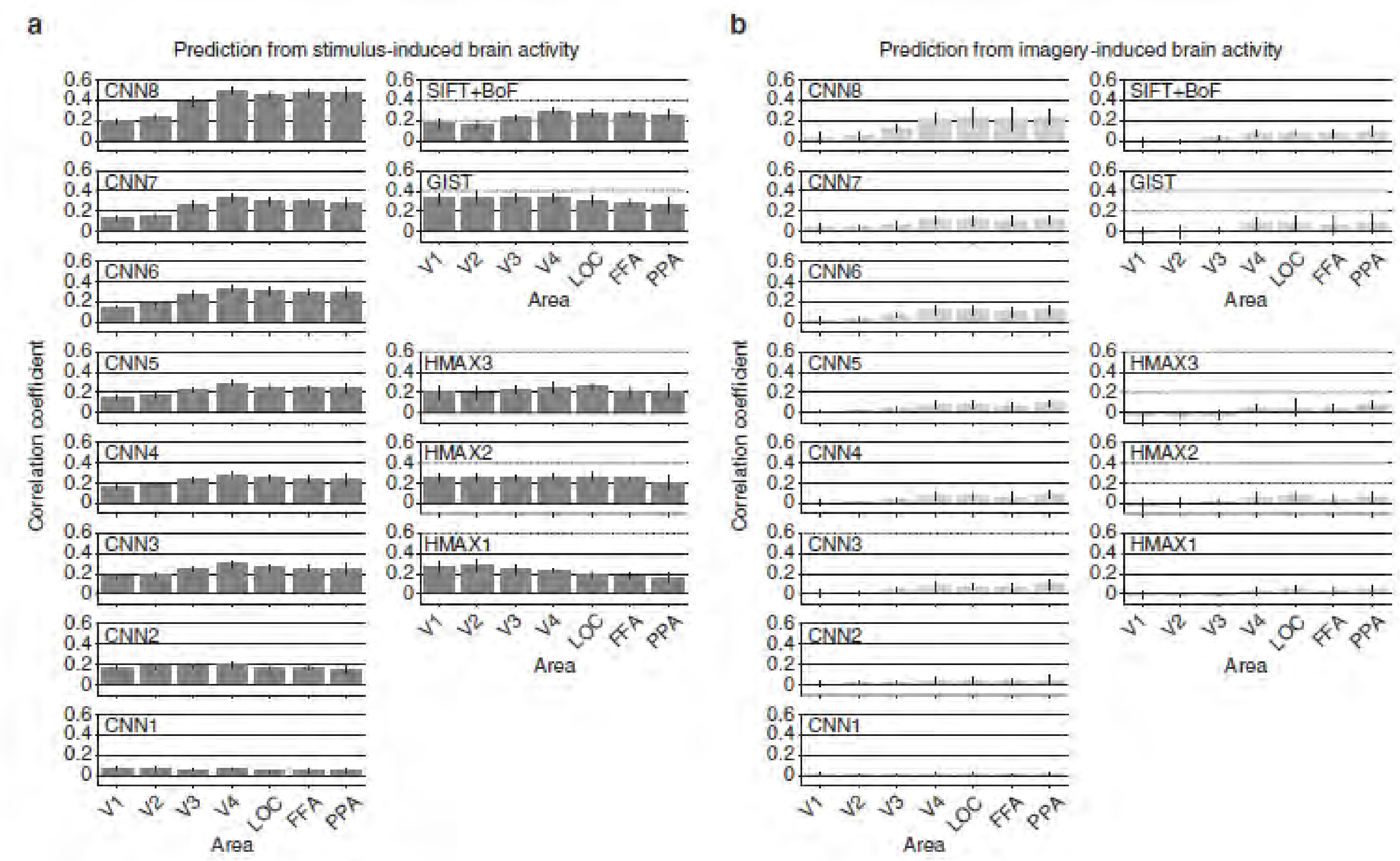
Decoding with DNN features





Horikawa & Kamitani, 2017

Decoding with DNN features



Horikawa & Kamitani, 2017

DNNs and brain decoding

- DNNs can be used to decode mental content of images
- What about the mental contents of language?
- Can current NLP models help?

Decoding linguistic information

- Natural language processing (NLP): word embeddings
  - Fairly low dimensional (e.g., 300 or 500 dimensions)
  - A word/sentence is represented by a vector in this space.
  - Vector operations make sense
  - Word/sentence embedding or "latent space"

vector $\vec{x}$ defined as:	Example 1	Example 2
$\vec{x}$ = Paris – France	Italy + $\vec{x}$ = Rome	Japan + $\vec{x}$ = Tokyo
$\vec{x}$ = bigger – big	cold + $\vec{x}$ = colder	quick + $\vec{x}$ = quicker
$\vec{x}$ = scientist – Einstein	Messi + $\vec{x}$ = midfielder	Mozart + $\vec{x}$ = violinist
$\vec{x}$ = Cu – copper	zinc + $\vec{x}$ = Zn	gold + $\vec{x}$ = Au
$\vec{x}$ = sushi – Japan	Germany + $\vec{x}$ = bratwurst	USA + $\vec{x}$ = pizza

Mikolov, T. *et al.* arXiv:1310.4546 (2013)

Decoding linguistic information

Experiment 1:

Bird

1. The bird flew around the cage.
2. The nest was just big enough for the bird.
3. The only bird she can see is the parrot.
4. The bird poked its head out of the hatch.
5. The bird holds the worm in its beak.
6. The bird preened itself for mating.

 . . . 

Nest, Flock, Bird, Mating, Beak, Winged

Wash

1. To make the counter sterile, wash it.
2. The dishwasher can wash all the dishes.
3. He likes to wash himself with bar soap.
4. She felt clean after she could wash herself.
5. You have to wash your laundry beforehand.
6. The maid was asked to wash the floor.

 . . . 

Clean, Shower, Sink, Wash, Soap, Laundry

Unaware

1. She was unaware of how oblivious he really was.
2. She was unaware of her status.
3. Unprejudiced and unaware, she went full throttle.
4. Unaware of current issues, he is a terrible candidate.
5. He was unaware of how uninterested she was.
6. He was unaware of the gravity of the situation.

 . . . 

Unprepared, Unprotected, Unaware, Unwilling, Inexperienced, Unconcerned

Experiment 2:

Musical instruments (clarinet)

A clarinet is a woodwind musical instrument.
It is a long black tube with a flare at the bottom.
The player chooses notes by pressing keys and holes.
The clarinet is used both in jazz and classical music.

Musical instruments (accordion)

An accordion is a portable musical instrument with two keyboards. One keyboard is used for individual notes, the other for chords. Accordions produce sound with bellow that blow air through reeds. An accordionist plays both keyboards while opening and closing the bellows.

Musical instruments (piano)

The piano is a popular musical instrument played by means of a keyboard. Pressing a piano key causes a felt-tipped hammer to hit a vibrating steel string. The piano has an enormous note range, and pedals to change the sound quality. The piano repertoire is large, and famous pianists can give solo concerts.

Experiment 3:

Skiing (passage 1)

I hesitantly skied down the steep trail that my buddies convinced me to try. I made a bad turn, and I found myself tumbling down. I finally came to a stop at a flat part of the slope. My skis were nowhere to be found, and my poles were lodged in a snow drift up the hill.

Skiing (passage 2)

A major strength of professional skiers is how they use ski poles. Proper use of ski poles improves their balance and adds flair to their skiing. It minimizes the need for upper body movements to regain lost balance while skiiing.

Skiing (passage 3)

New ski designs and stiffer boots let skiers turn more quickly. But faster and tighter turns increase the twisting force on the legs. This has led to more injuries, particularly to ligaments in the skier's knee.

Gambling (passage 1)

When I decided to start playing cards, things went from bad to worse. Gambling was something I had to do, and I had already spent close to $10,000 doing it. My friends were sick of watching me gamble my savings away. The hardest part was the horror of leaving a casino after losing money I did not have.

Gambling (passage 2)

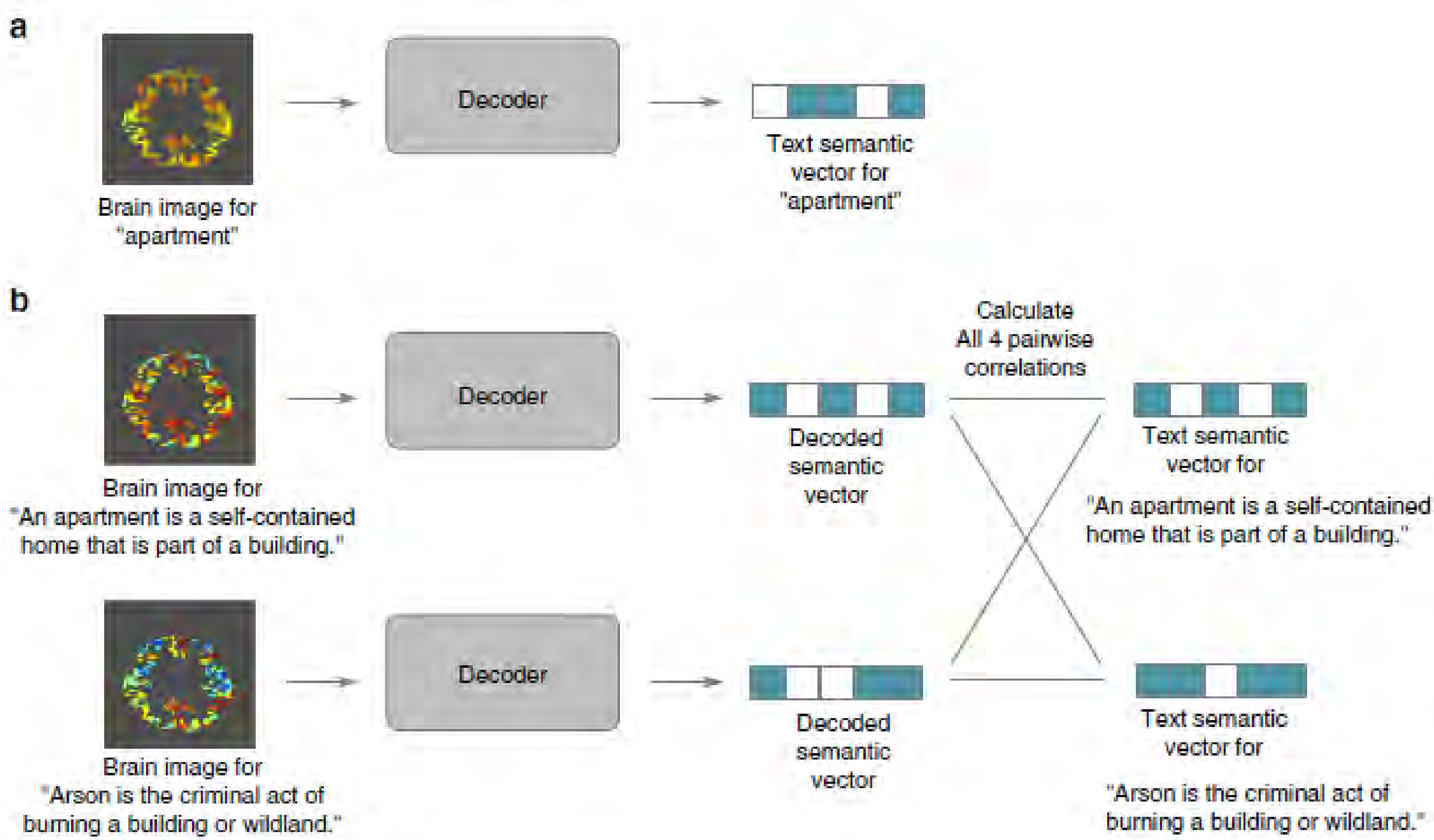
Good data on the social and economic effects of legalized gambling are hard to come by. Some studies indicate that having a casino nearby makes gambling problems more likely. Gambling may also be associated with personal bankruptcies and marriage problems.

Gambling (passage 3)

Over the past generation, there has been a dramatic expansion of legalized gambling. Most states have instituted lotteries, and many have casinos as well. Gambling has become a very big but controversial business.

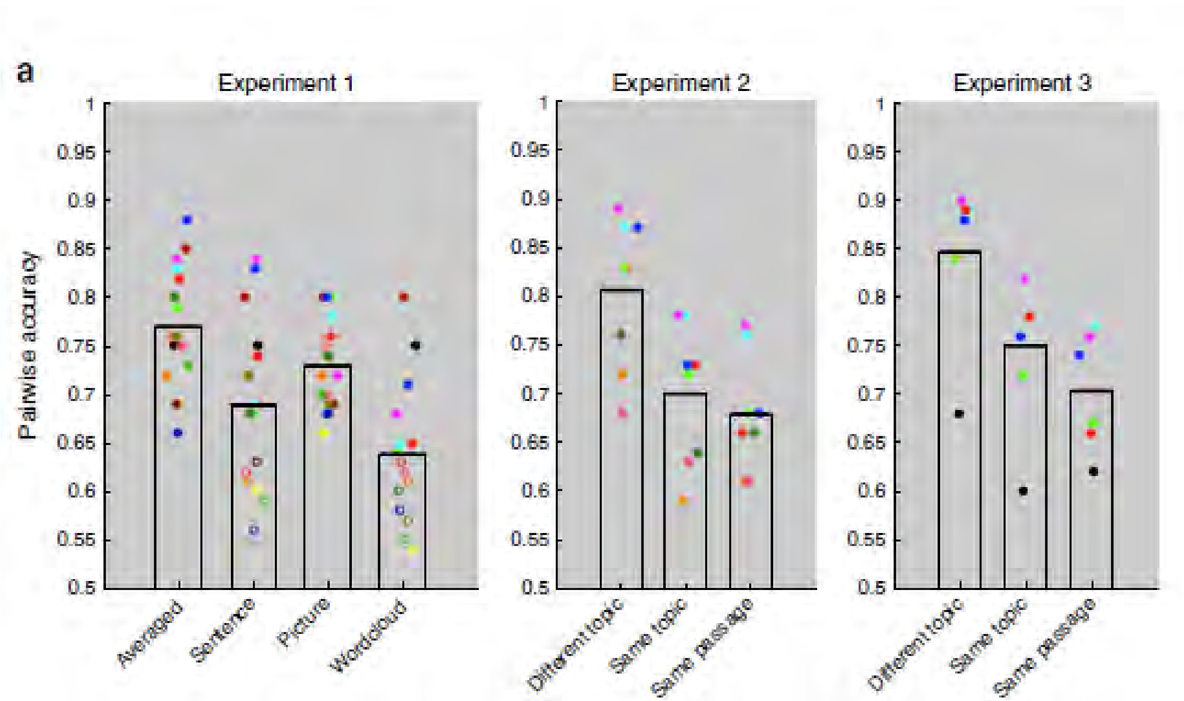
Pereira et al., 2018

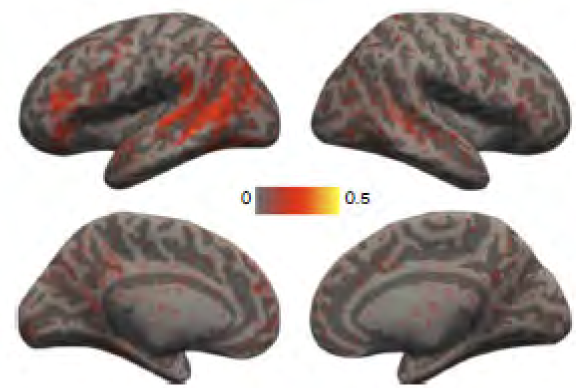
Decoding linguistic information



Pereira et al., 2018

Decoding linguistic information





Pereira et al., 2018

Stimulus reconstruction

- So far:
  - Decoding with features and SVMs
  - Decoding using the DNN feature space
  - Decoding => guess the stimulus (image/category/ sentence)

- Stimulus Reconstruction
  - Reconstruct the face or image a subject is looking at.
  - A whole new level of « decoding ».
  - Generative models open up a range of possibilities.

Face latent spaces

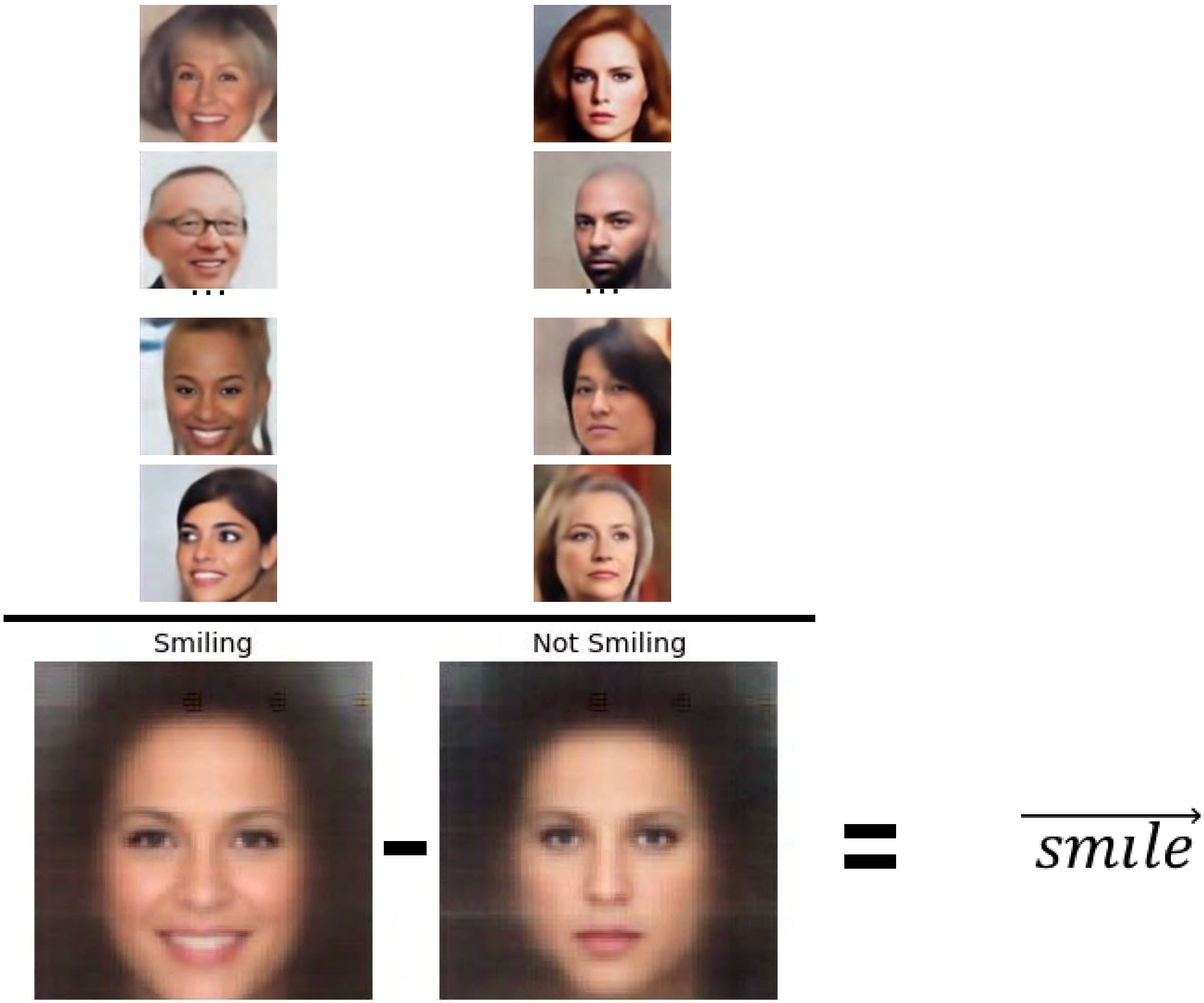
Computer Vision:

- Example: a GAN trained on (200K) celebrity faces
- Creates a latent space, e.g., a 500 or 1000 dimensional space
- A point/vector in this space corresponds to a face
- GAN: generative model ➔ generate a face from a vector
- Perform operations on these vectors and look at the faces that are generated
- Vector operations make sense?

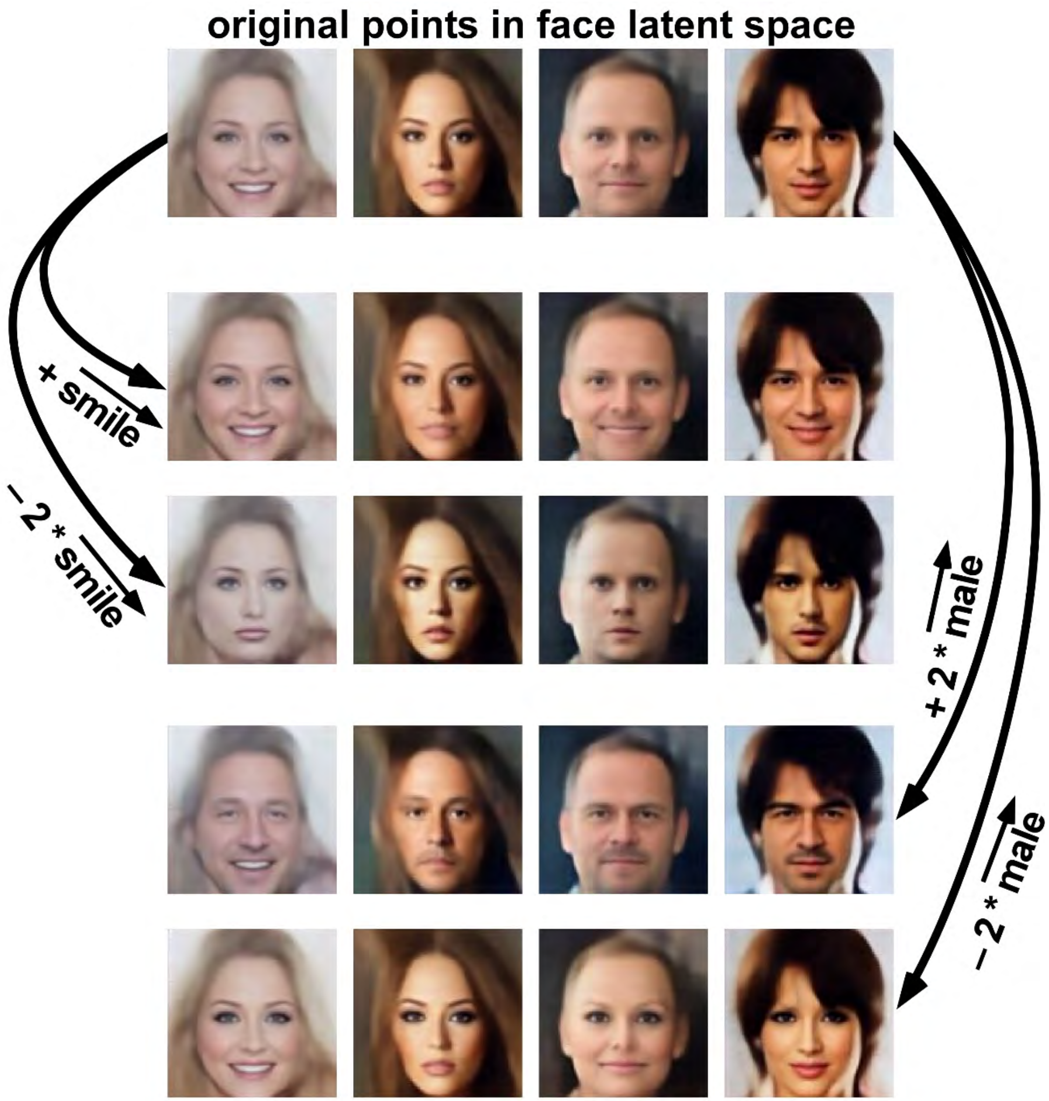
Face latent spaces

White,T. (2016) *arXiv:1609.04468*.

Face latent spaces



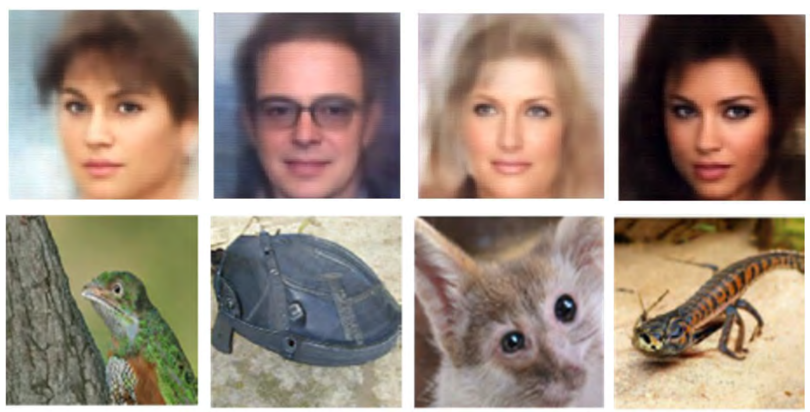
Face latent spaces

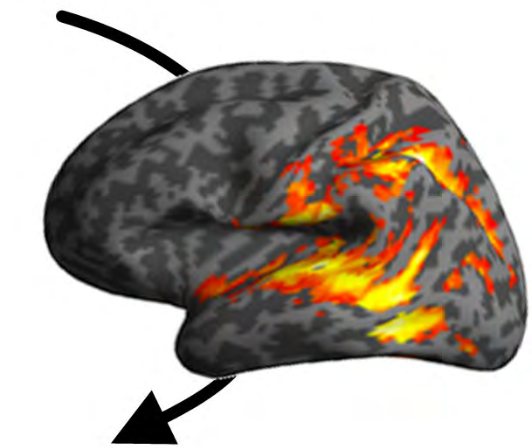


Decoding faces

images seen by subject in MRI scanner

images reconstructed from fMRI signals





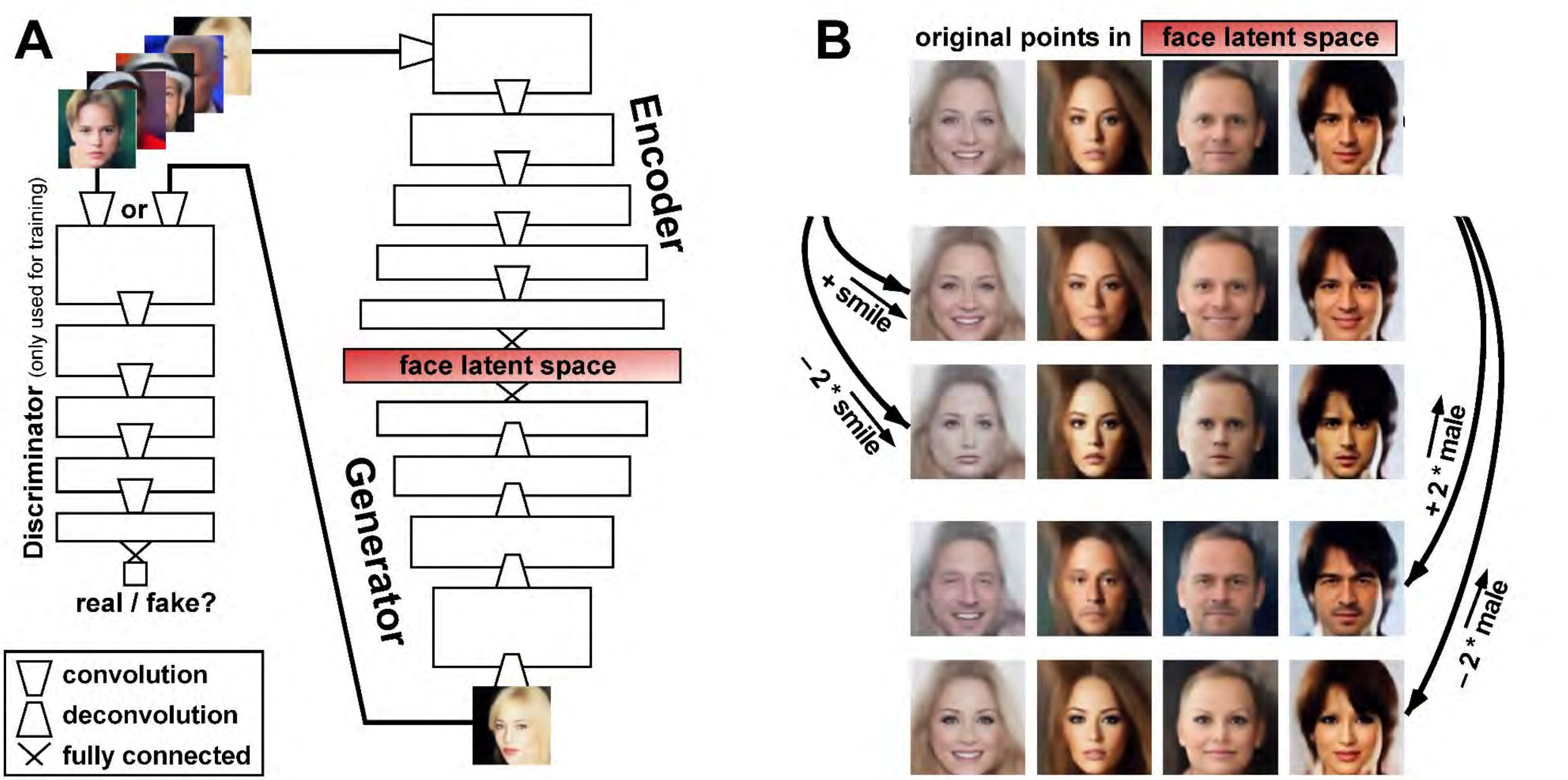
VanRullen & Reddy (2019) *Communications Biology*
Mozafari, Reddy & VanRullen (2020) *IJCNN'20*

VAE-GAN model

The network is trained on a celebrity dataset (200,000 images).

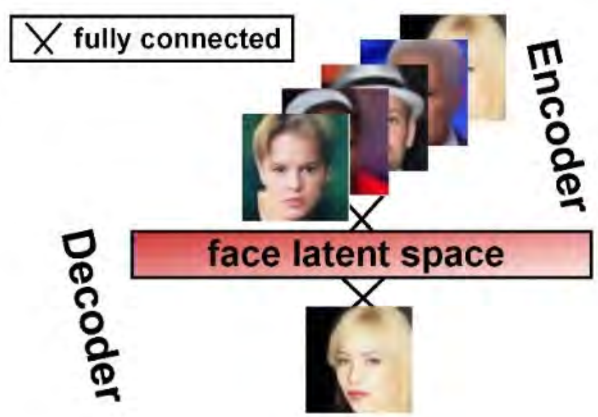
The encoder defines a "latent space" of 1024 dimensions.

After training, the weights are frozen and the discriminator network is dropped.



Larsen, A. B. L., Sønderby, S. K., Larochelle, H. & Winther, O. *ICML 2016*.

PCA as a comparison model

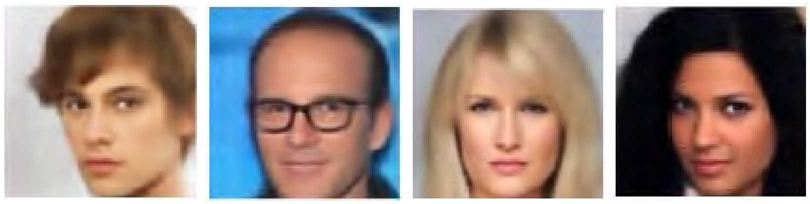


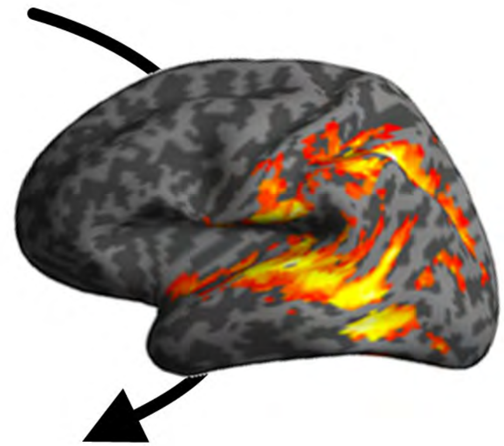
Faces are encoded into a latent space of 1024 principal components

➔ Previous state-of-the-art for face reconstruction from fMRI (Cowen et al., 2014; Lee et al., 2016.).

Face decoding via VAE-GAN latent space

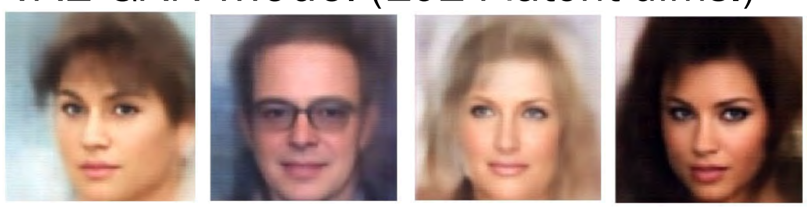
images seen by subject in MRI scanner



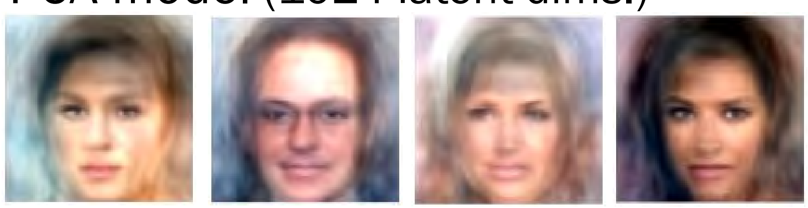


images reconstructed from fMRI signals

VAE-GAN model (1024 latent dims.)



PCA model (1024 latent dims.)



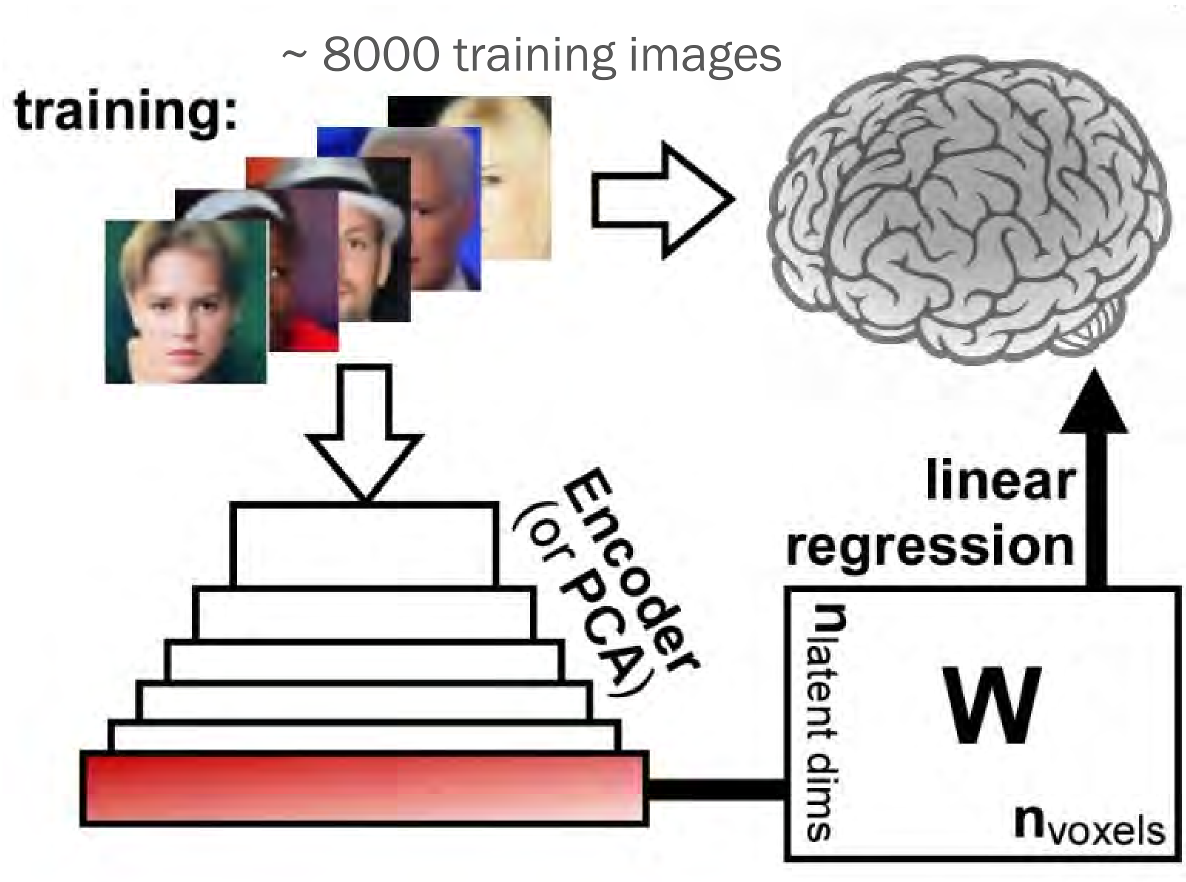
Cowen et al., 2014; Lee et al., 2016.

- 4 subjects
- ~8000 faces/subject
- 16 hours of scanning/subject

➔ Freely available on OpenNeuro

VanRullen & Reddy, 2019. *Communications Biology*

Training the decoder



$$Y = XW \quad \Leftrightarrow \quad X^TY = X^TXW$$

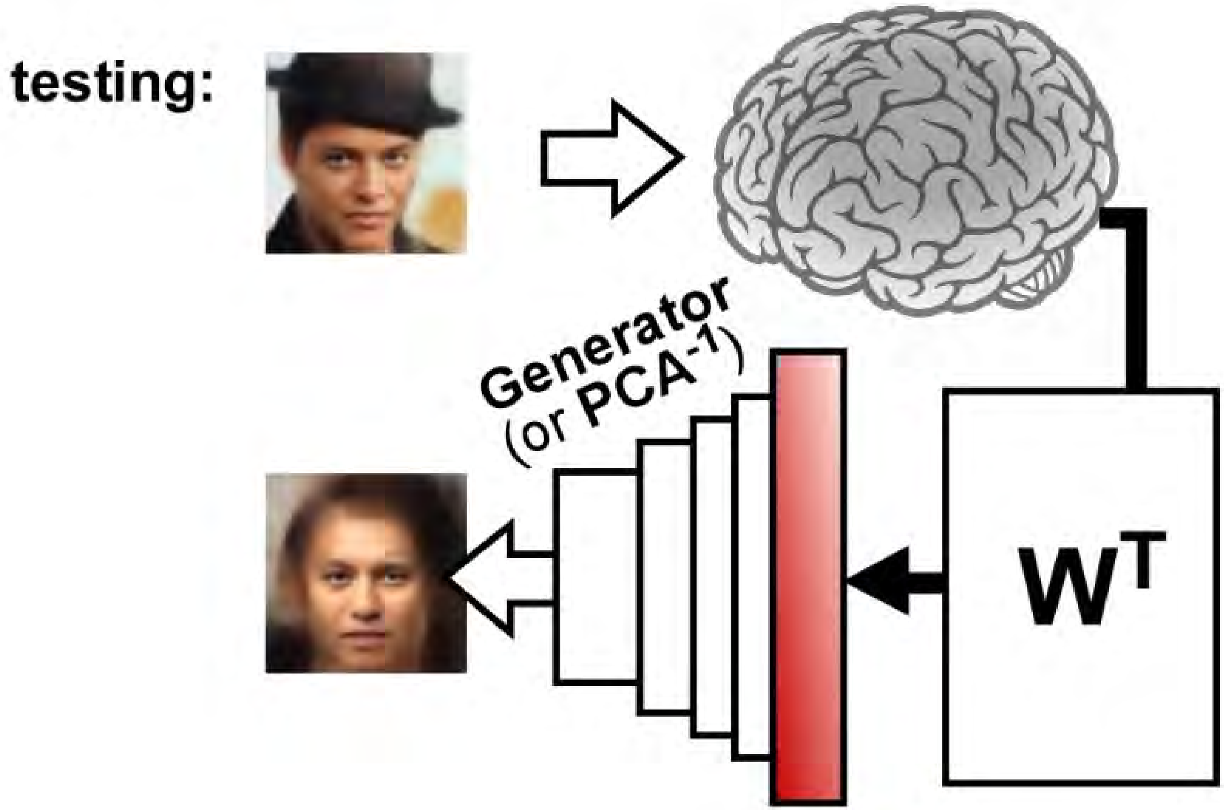
Y: (8000, n_{voxels}) X: (8000,1024) W: (1024, n_{voxels})

$$\boxed{W = (X^TX)^{-1}X^TY}$$

VanRullen & Reddy, 2019. *Communications Biology*

Testing the decoder

20 test images (x 45 repeats)



$$Y = XW$$

$(20, n_{voxels})$ $(20, 1024)$ $(1024, n_{voxels})$

$$\Leftrightarrow$$

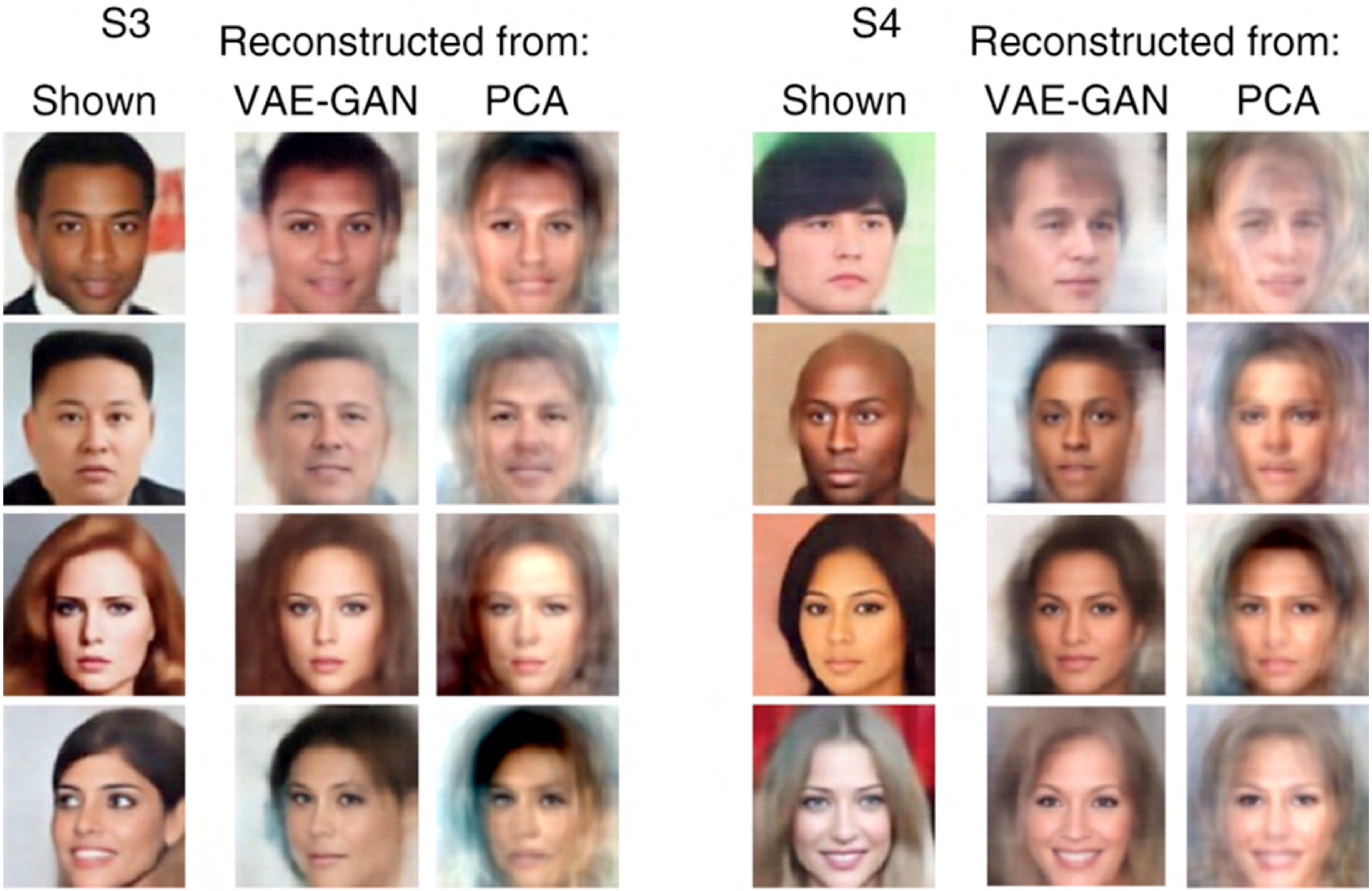
$$YW^T = XWW^T$$

$$\boxed{X = YW^T(WW^T)^{-1}}$$

VanRullen & Reddy, 2019. *Communications Biology*

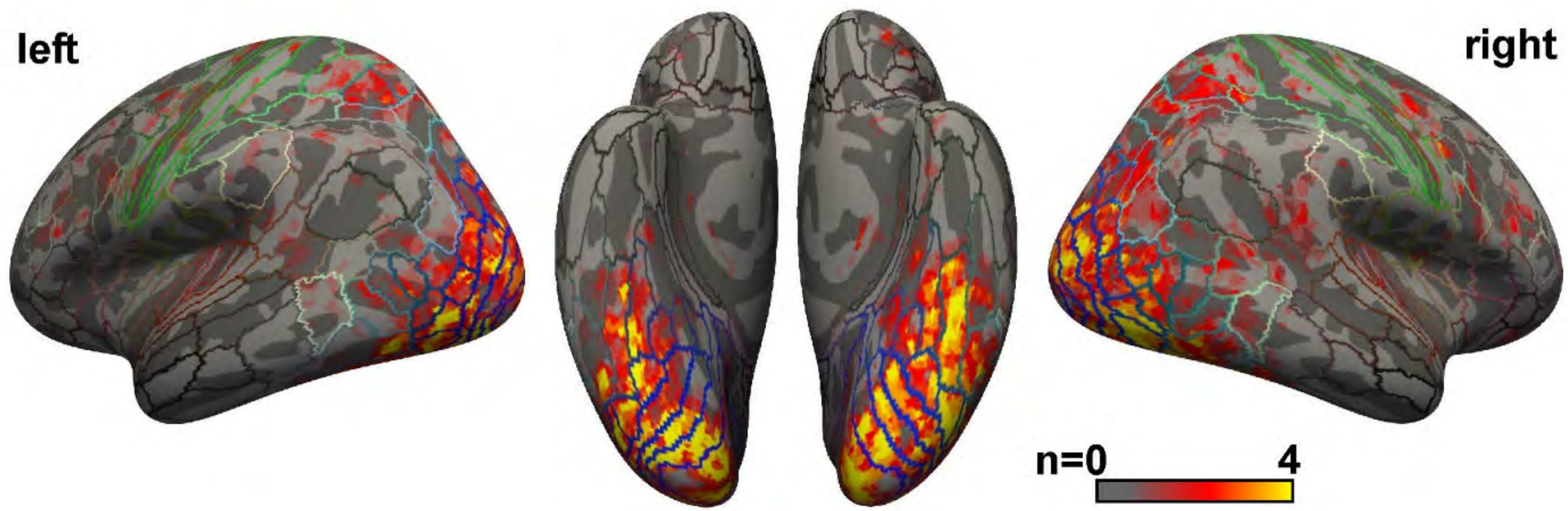
Face Reconstructions

Face Reconstructions



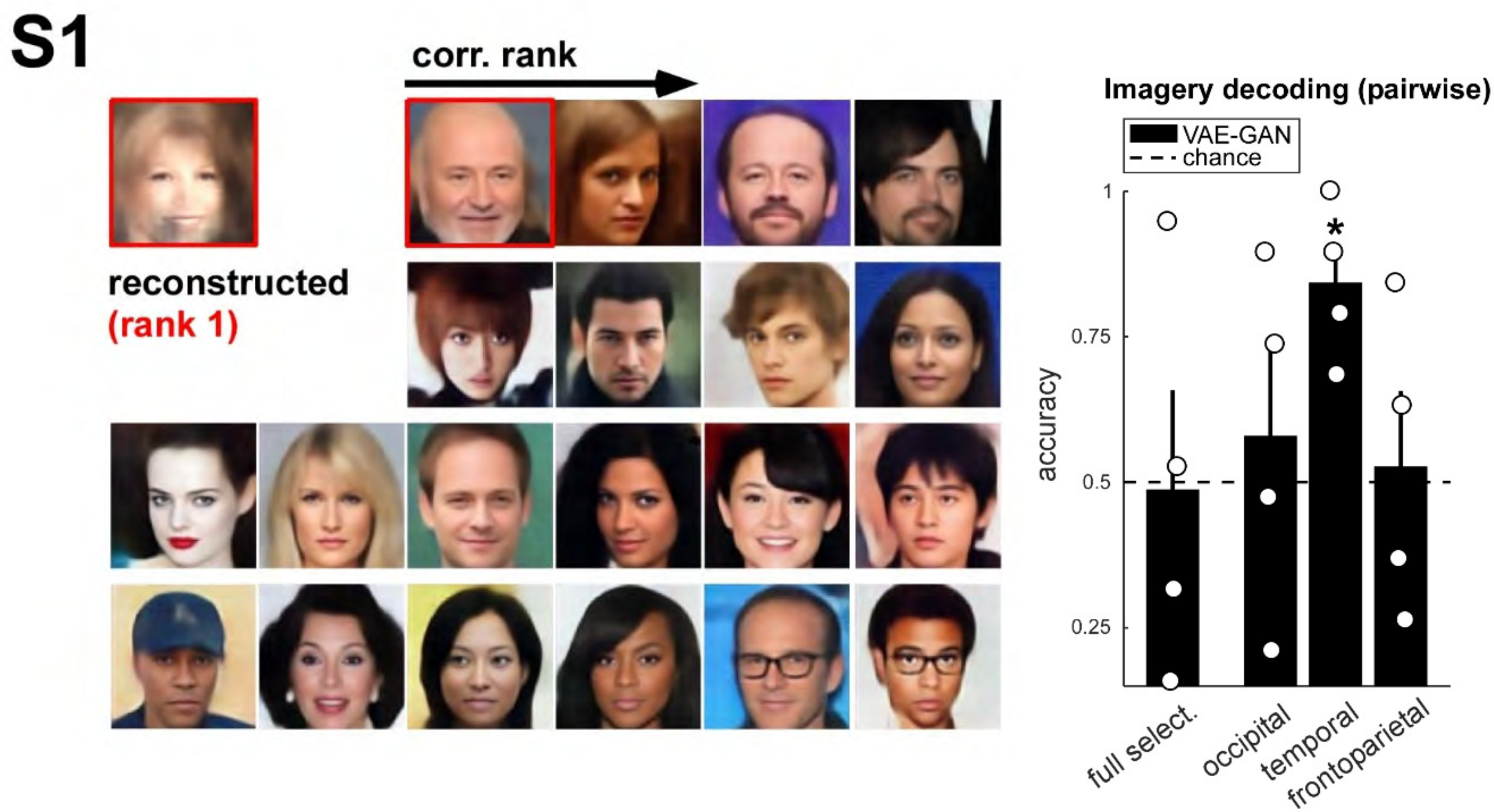
VanRullen & Reddy, 2019. *Communications Biology*

Which brain regions are involved?



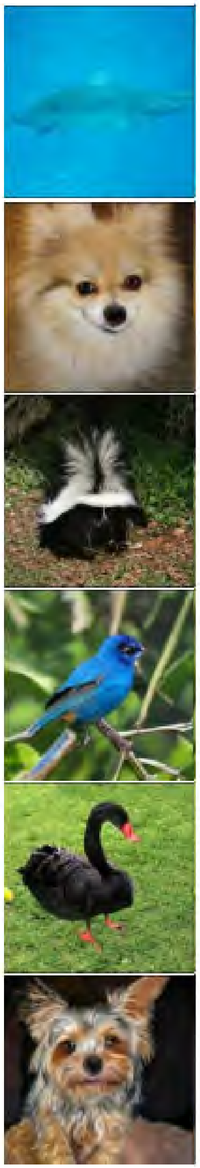
VanRullen & Reddy, 2019. *Communications Biology*

Decoding mental imagery



VanRullen & Reddy, 2019. *Communications Biology*

A GAN for all visual scenes?



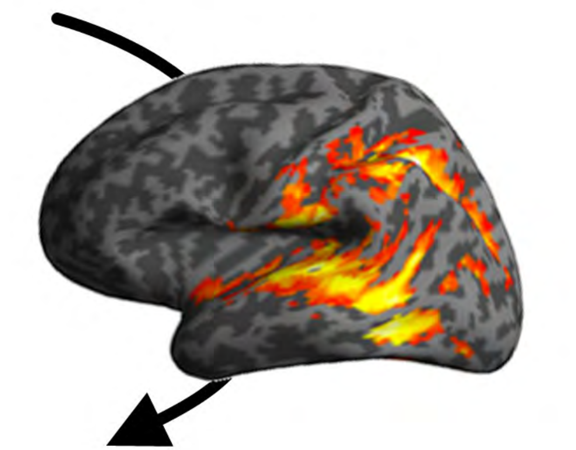
BigGAN Brock, Donahue & Simonyan, ICLR (2019)

- Can generate any image category (among 1000 ImageNet classes)
- These animals do not exist!
- They are “just” points in a latent space

fMRI decoding via BigBIGAN latent space

Mozafari, Reddy & VanRullen, 2020. *IJCNN'20*

images seen by subject in MRI scanner



images reconstructed from fMRI signals

BigBiGAN model (120 latent dims.)

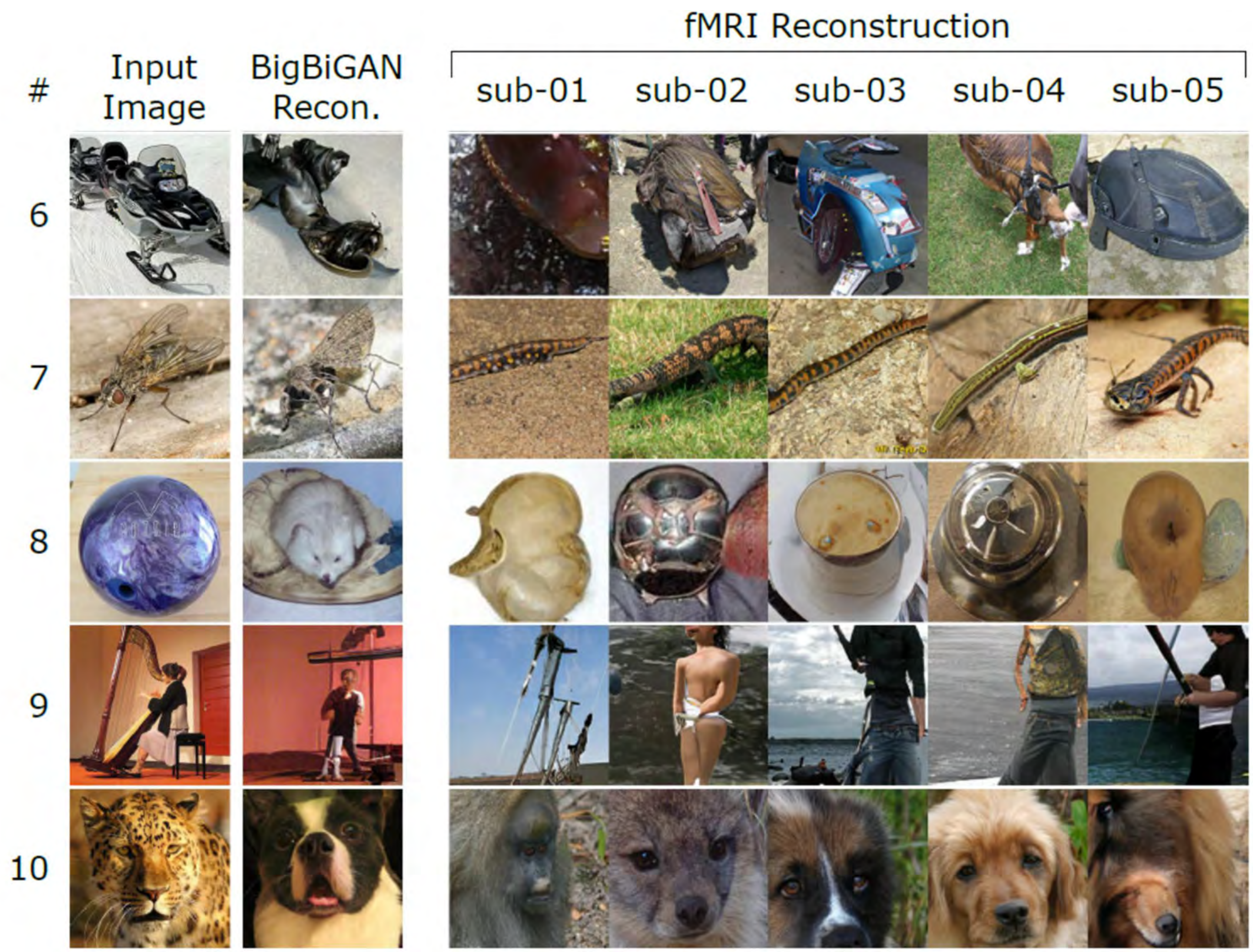
PCA model (120 latent dims.)

Cowen et al., 2014; Lee et al., 2016.

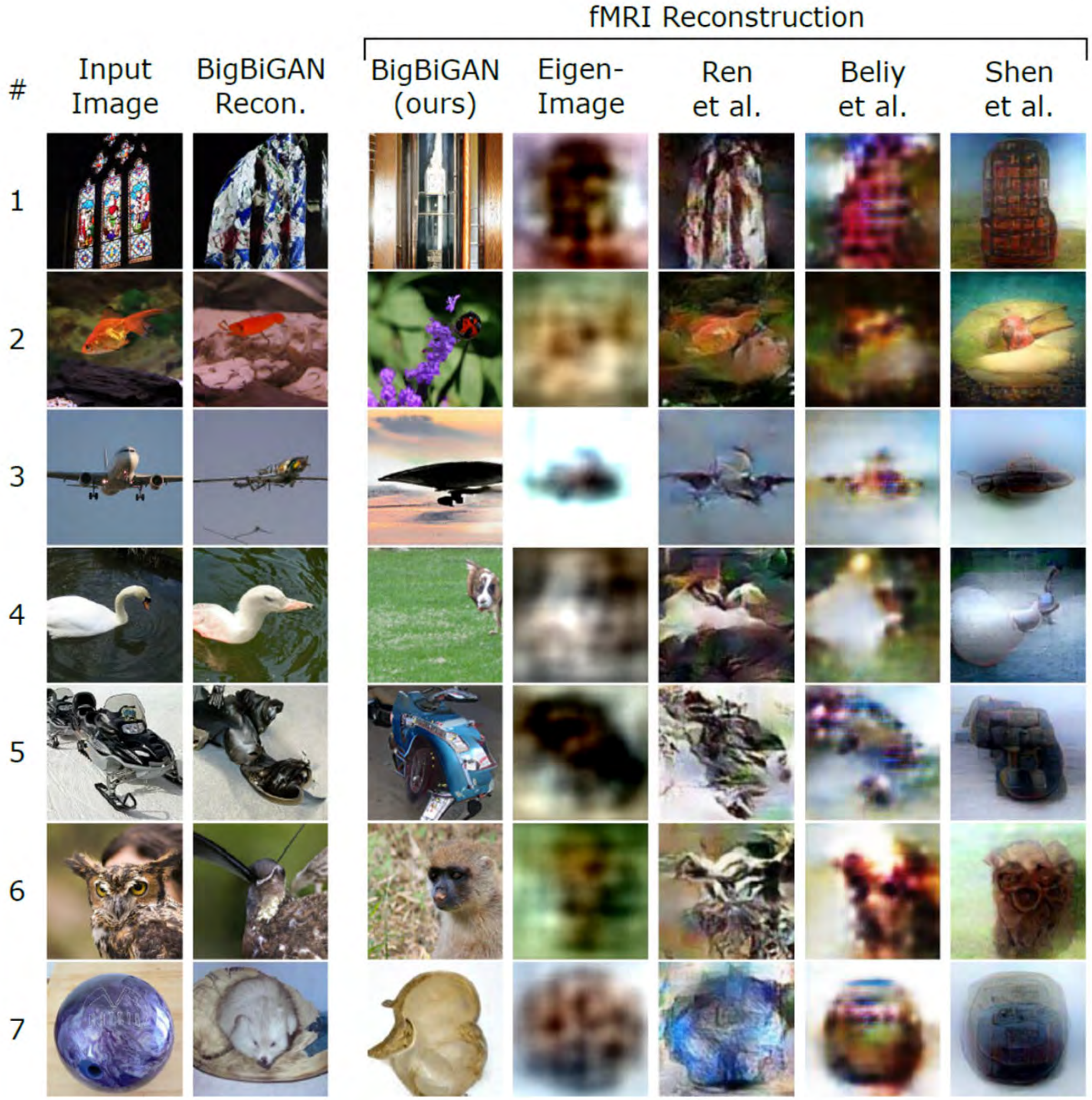
- Data from Horikawa & Kamitani, 2017 (available on OpenNeuro)
- 5 subjects
- ~1200 images/subject

fMRI decoding via BigBIGAN latent space

Mozafari, Reddy & VanRullen, 2020. *IJCNN'20*



fMRI decoding via BigBIGAN latent space



Summary

- Exploit patterns of brain activation to decode stimulus information.
- Classification methods (e.g., SVM) that allow us to guess the stimulus (restricted to training stimuli).
- Using DNN feature space for decoding allows us to generalize to new images/categories/sentences.
- Generative models open up a new range of possibilities: reconstruction of what the subject was looking at.

Questions?