University Medicine Essen
Institute for Artificial Intelligence in Medicine

# Disentanglement in Neuroimage Analysis:

## a quick overview of theories and applications

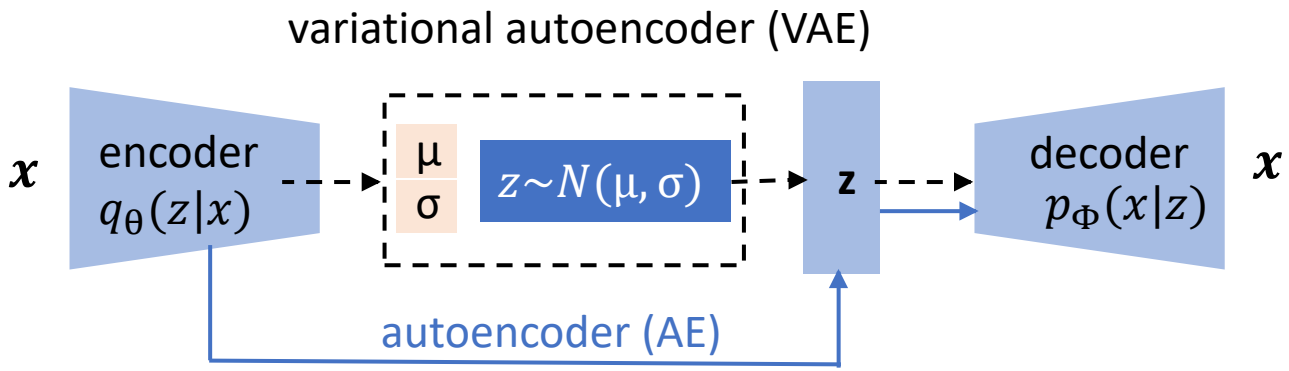
Jianning Li

Institute for Artificial Intelligence in Medicine (IKIM)
University Hospital Essen (AöR)

# Outline

- Disentanglement
  - variational auto-encoder (VAE)
  - human face generation
- Brain aging
- Longitudinal brain imaging of neurological disorders
  - Alzheimer‘s disease (the most common type of dementia)
  - Brain tumor
- Understanding how the biological brain works
  - How the brain perceives faces (vision, face recognition)
  - How the brain perceives languages (speech comprehension)
- Opportunities and Challenges

## Disentangled Representation Learning




$$E_{z\sim q_\theta(z|x)}\left[\log p_\phi(x|z)\right] - D_{KL}(q_\theta(z|x)||p(z)) \quad , p(z) = N(0,1)$$

reconstruction term

latent space regularization (KLD loss)

- Posterior $q_\theta(z|x) \approx N(0,1)\ , z \in R^d$
- The regularization term (KLD loss) forces the covariance matrix of the latent variables ($z$) to be diagonal
- The dimensions of the a latent variable ($z$) are independent

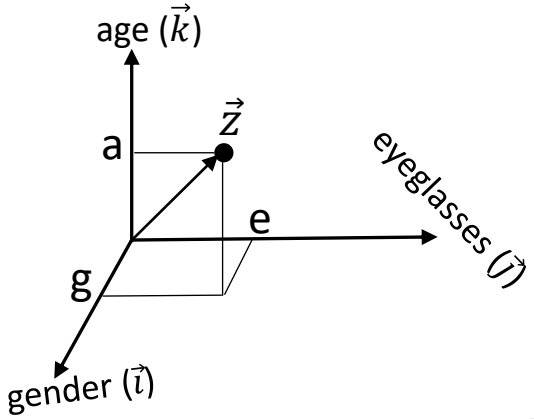

- Independence is a precondition of the representations being **disentangled**
- VAE is a popular choice for disentangled representation learning

face factors: age, gender, wear eyeglasses, pose, expressions (cry, smile, etc)…

[Shen, Y. et al. *TPAMI*, 44 (4), pp 2004 - 2018]

disentangled


$\vec{z}$ | a | e | g




$$\vec{z} = g * \vec{i} + e * \vec{j} + a * \vec{k}$$

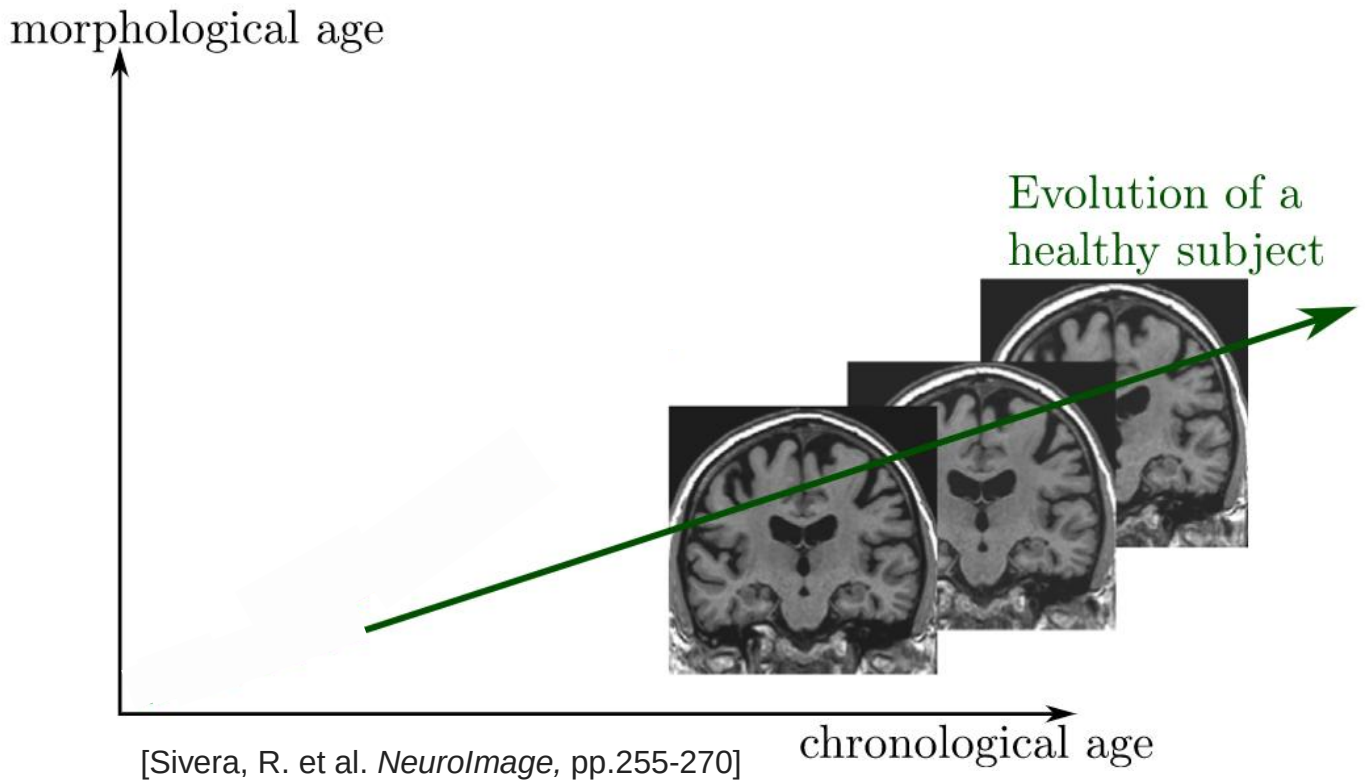
entangled


$\vec{z}$ | a+g | … | …

$$g = \vec{z} \cdot \vec{i}$$
$$e = \vec{z} \cdot \vec{j}$$
$$a = \vec{z} \cdot \vec{k}$$

- Factors must be inherently independent/uncorrelated (such as ‘age‘ and ‘gender‘) in order to be disentangled
- Such a disentangled representation allows us to generate new faces that are different from the original face regarding only one factor
- Each disentangled factor is a basis of the latent space, and all latent variables can be expressed as a linear combination of the three bases. An inner product between the latent representations $\vec{z}$ and a basis ($\vec{i}, \vec{j}$ or $\vec{k}$) yields the corresponding factor

***goal:*** *to model the morphological brain changes induced by normal brain aging*
***method:*** *by disentangling the "age" factor from brain MRIs*




[Sivera, R. et al. *NeuroImage,* pp.255-270]

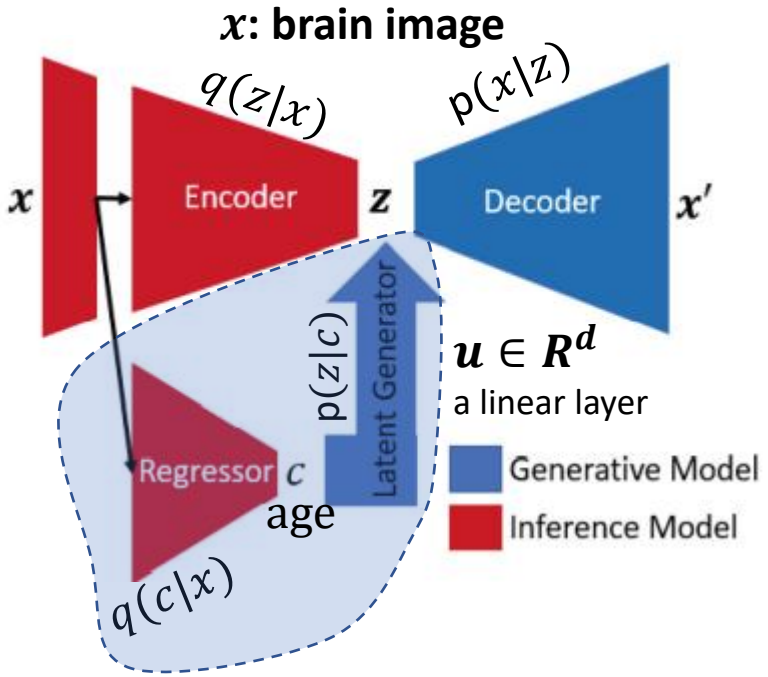
medical images (e.g., MRIs)

- identity factors: **age**, gender, etc of a patient
- medical factors: disease/pathology (e.g., Alzheimer's disease)

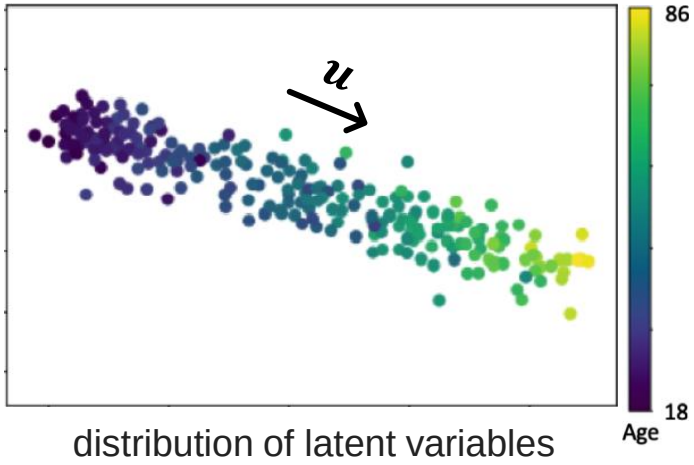
VAE: $z \sim q(z|x), x' \sim p(x|z)$

Age: $c \sim q(c|x), z \sim p(z|c)$

$$L(x) := \log q(c|x) + E_{q(Z|X)}[\log p(x|z)] - E_{q(c|x)}[D_{KL}(q(z|x)||p(z|c)]$$

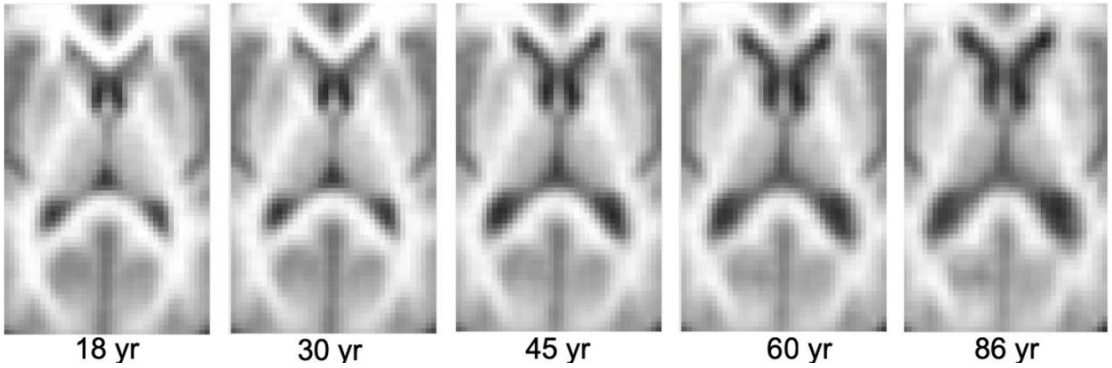



- latent representations $z$ are conditioned on age via a linear layer $u$
- $z = cu \leftrightarrow u$ is the basis of age (age c is a scalar)
- traversing along the $u$ direction yields age-specific latent variables and brain MRI reconstructions




distribution of latent variables

morphological brain changes with aging $u$

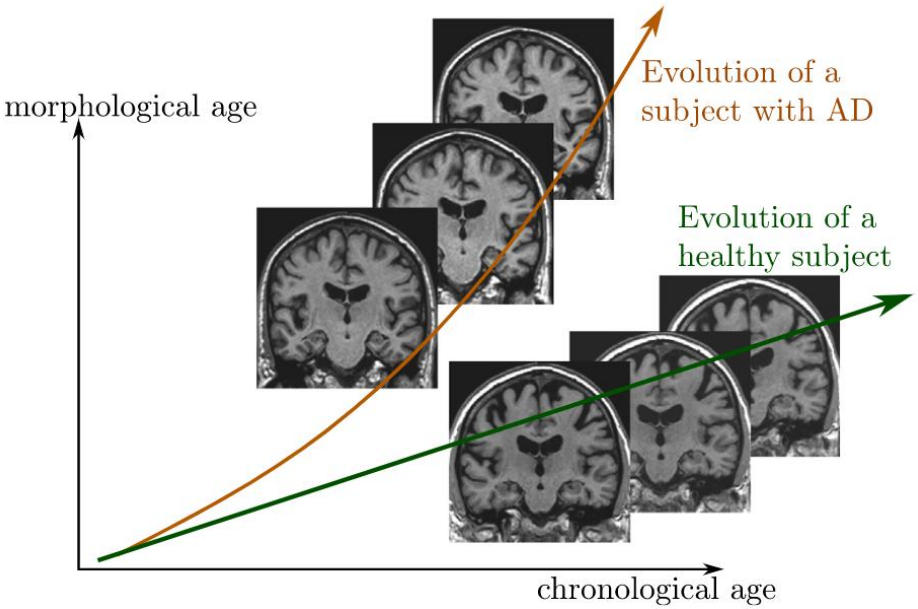



In the generated MRIs, the **ventricle expands with aging**, which is consistent with current clinical understanding of brain development

[Zhao, Q. et al. MICCAI 2019, pp. 823-831]

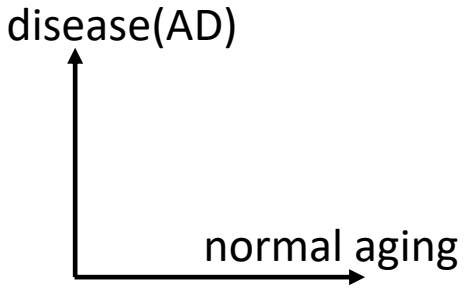
longitudinal MRIs are commonly used to track the progression of the neurological diseases:

- Alzheimer‘s disease (the most common type of dementia)
- Brain tumor




[Sivera, R. et al. *NeuroImage,* pp.255-270]

- *both normal aging and Alzheimer‘s disease (AD) can cause morphological brain changes*
- *we want to separate the AD effect on brain morphology change from normal aging*
- *it is assumed that nornal aging and AD are assumed to be two independent factors*

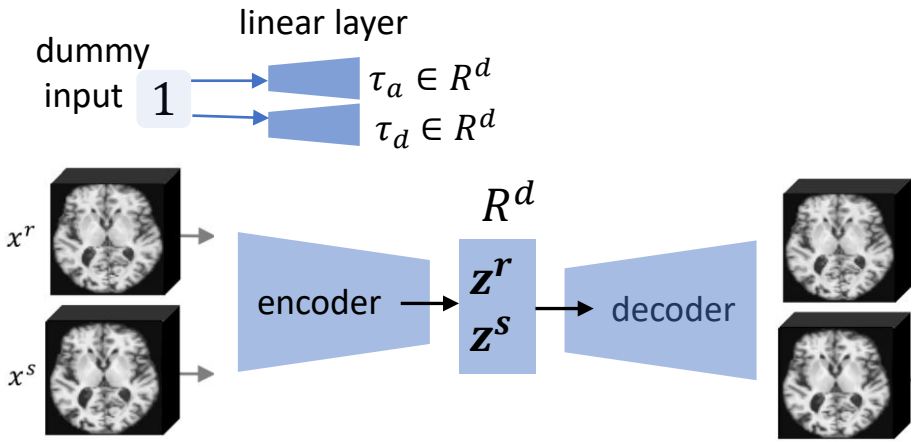



- *the two factors (normal aging and AD) can be disentangled*

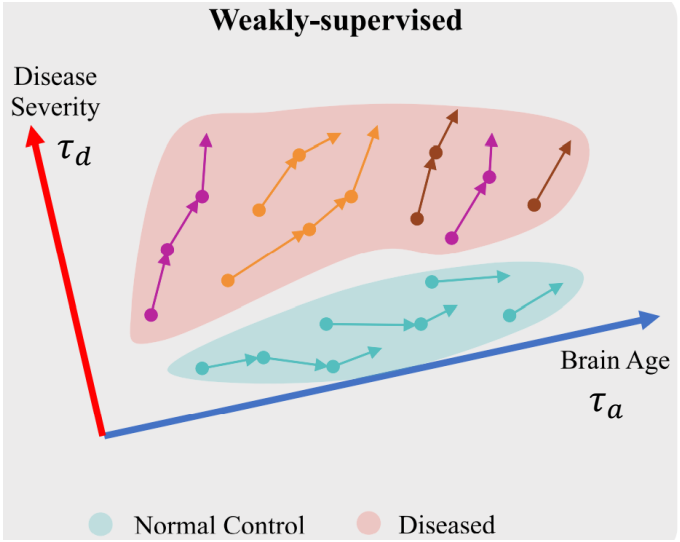
***for brain tumor:***

***goal:*** *to model the treatment response of a single cancer drug*
***issues:*** *the effects of different drugs on tumor response are NOT independent. if patients receive a combination of drugs for treatment, the effect of a single drug cannot be singled out*

*[Ouyang, J. et al. IEEE TMI, 41(10) pp. 2558 - 2569]: learn the two bases (i.e., $\tau_a, \tau_d$) directly*

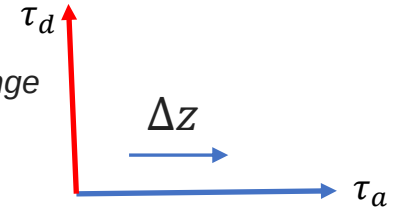



input: a pair of MRIs taken from the same subject but at different time points, and a dummy variable (1)



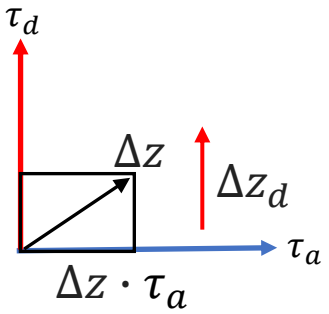

## *for normal cases (neurologically healthy subjects)*

- *aging is the only time-dependent factor for brain morphological change*
- $\Delta z = z^s - z^r // \tau_a \rightarrow \max \cos(\theta < \Delta z, \tau_a >)$




## *for diseased (AD) cases*

- *both normal aging and AD affect the brain morphology*
- *aging and AD are the only two time-dependent factors affecting the brain morphological change*
- $\Delta z_d = \Delta z - \Delta z_a // \tau_d \rightarrow \max \cos(\theta < \Delta z_d, \tau_d >)$




$$\Delta z = \Delta z_a + \Delta z_d$$
$$\Delta z_a = (\Delta z \cdot \tau_a)\, \tau_a$$

## to simulate the aging effect

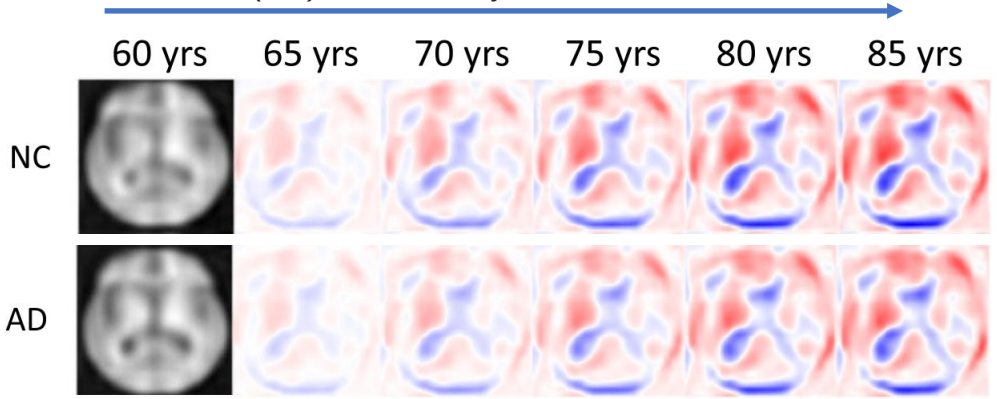
$$\hat{z}^C := \hat{\varphi}_a \tau_a + \frac{1}{N^C} \sum_{i=1}^{N^C} (z^i - \varphi_a^i \tau_a)$$

$$\varphi_a^i = z^i \cdot \tau_a$$

$N^C$ *number of normal controls (NC)*

$\hat{\varphi}_a$ *age (a scalar) to be simulated*

*simulate the effect of normal aging on brain morphology change for normal (NC) and AD subjects*




## to simulate the AD effect

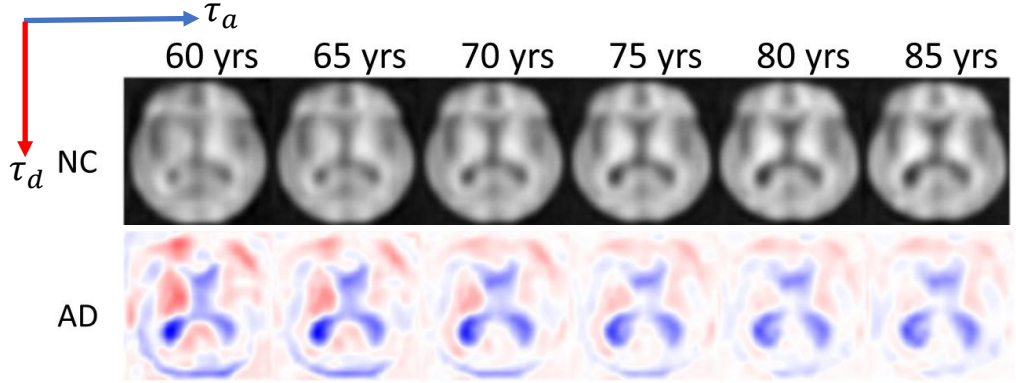
$$\hat{z}^{D_k} = \hat{\varphi}_a \tau_a + \left( \frac{1}{N^{D_k}} \sum_{i=1}^{N^{D_k}} \varphi_d^i \right) \tau_d + \frac{1}{N^{D_k}} \sum_{i=1}^{N^{D_k}} \left( z^i - \varphi_a^i \tau_a - \varphi_d^i \tau_d \right)$$

$$\varphi_d^i = z^i \cdot \tau_d$$

$N^{Dk}$: *number of normal AD subjects*

$\hat{\varphi}_a$ *age (a scalar) to be simulated*

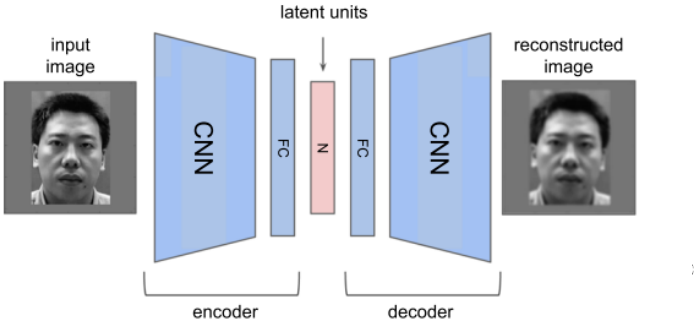
*simulate the effect of AD on brain morphology change (compared to normal subjects) given different ages*



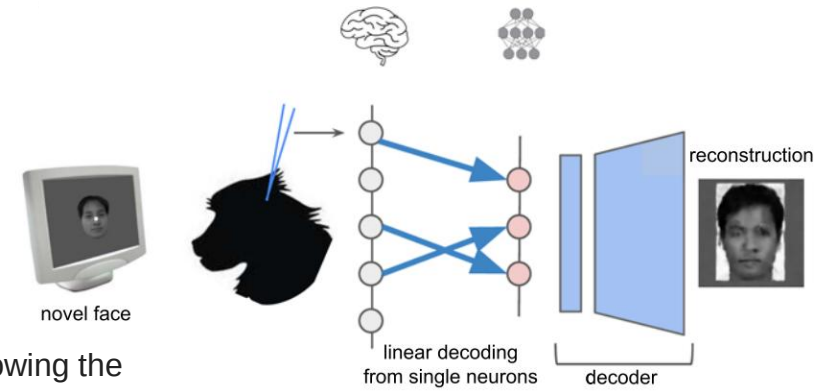

[Ouyang, J. et al. IEEE TMI, 41(10) pp. 2558 - 2569]

*different brain neurons respond to different facial factors (hair, gender, age, ethnicity, etc). In other words, the responses of brain neurons are disentangled in face perception.* [Higgins, I. et al. Nature communications, 12(1), pp.1-14]

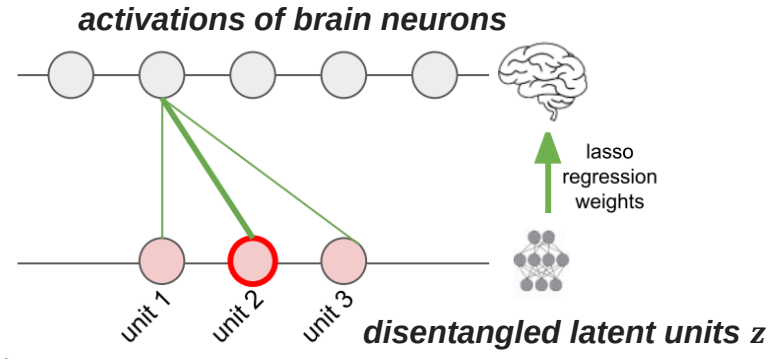
**(1)** use a VAE to learn disentangled facial representations $z$




**(3)** decode human face from the brain neuron activations of a primate




**(2)** record the brain neuron activations of a primate while showing the primate human faces. (green lines: regression weights)




Use lasso regression to predict neural activations from disentangled latent units $z$

each brain neuron has strong correlation (thickest green line) with only one disentangled latent unit (facial factor)

# *Face perception*

*which brain neuron responds to which facial factor*

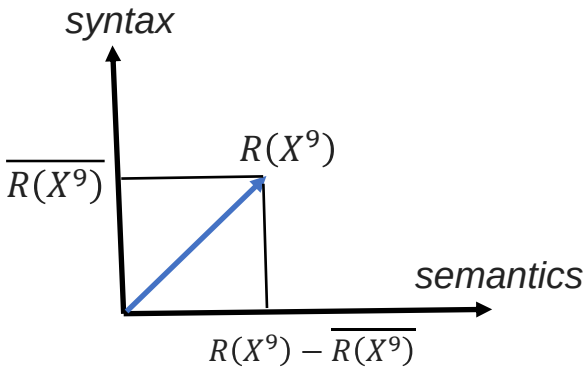
***Data acquisition: record the brain fMRI (≈ 4 hours) from 345 subjects while they are listening to stories (audio stimuli). Given GPT-2 (a pre-trained language transformer) and the subject the input (a sentence of M words)*** $w = (w_1, \ldots, w_M)$:

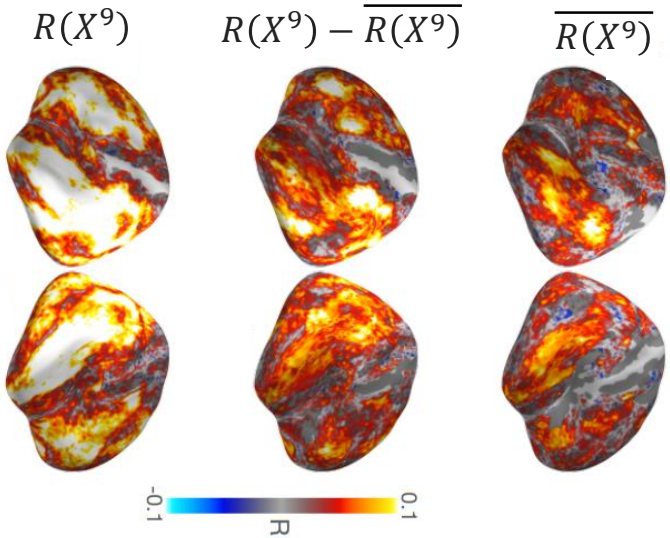
***(1)*** *Pretrained GPT-2 model*

- $R(X^9)$ *the activations extracted from the 9th layer of GPT-2*
- $\overline{R(X^9)}$ *the* ***syntactic*** *factor of w*
- $R(X^9) - \overline{R(X^9)}$ *the* ***semantic*** *factor of w*

***(2)*** *Map network activations to the corresponding brain fMRI recordings*

- *Mapping: ridge regression*
- *Mapping quality: Pearson correlation score (the brain score)*




*color code: brain scores*

*Latent representations learned by a language model such as* ***GPT-2*** *disentangle* ***syntax*** *(structure and grammar) and* ***semantics*** *(meaning and logic) of a sentence, which can be linearly mapped to brain activities [Caucheteux, C. et al. ICML 2021 pp. 1336-1348]*

# *Language/speech perception*

## Opportunities

- *Disentangled representation learning is advantangeous in modelling longitudinal data, which are common in neuroimage analysis*
- *Contrast to the black-box deep models, disentangled representations are human-interpretable, which offers a natural interface between deep learning and human domain knowledge*

## Challenges:

- *Data are harder to acquire, and ethical approval is more complicated (animal experiments, clinical trials on humans) than conventional medical data*
- *Biological factors are usually not strictly independent e.g, disease-age effect (neurological diseases cause accelerated aging), age-gender effect (gender plays a role in brain aging)*

[Coffey, C.E. *et al. Archives of neurology*, *55*(2), pp.169-179]
*[Király, A. et al. Brain imaging and behavior, 10(3), pp.901-910]*

University Medicine Essen
Institute for Artificial Intelligence in Medicine

# Disentanglement in Neuroimage Analysis:

## a quick overview of theories and applications

Jianning Li

Institute for Artificial Intelligence in Medicine (IKIM)
University Hospital Essen (AöR)