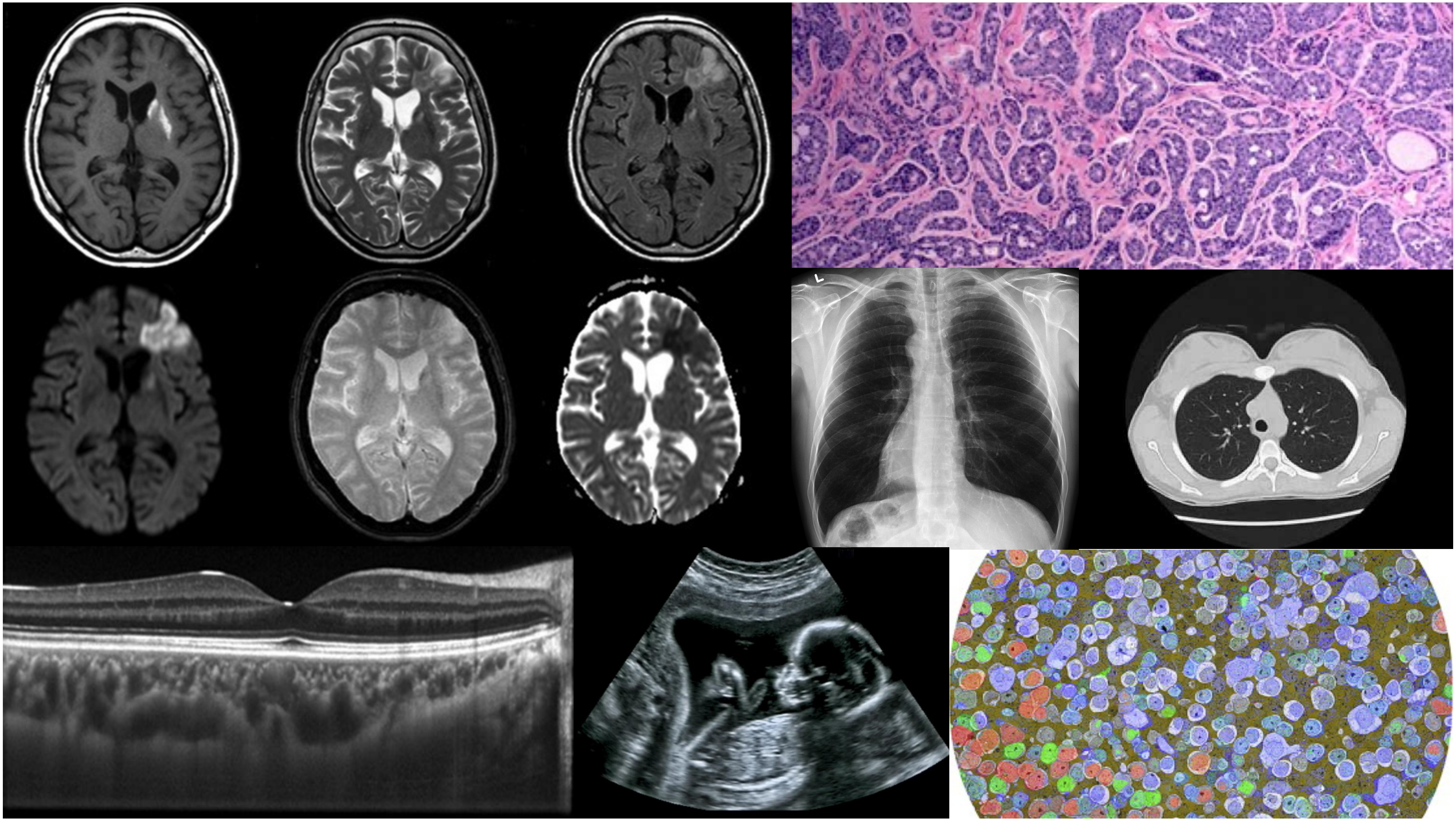
Machine Learning for Medical Image Analysis

Adrian V. Dalca

MIT CSAIL and
Massachusetts General Hospital, Harvard Medical School



Outline

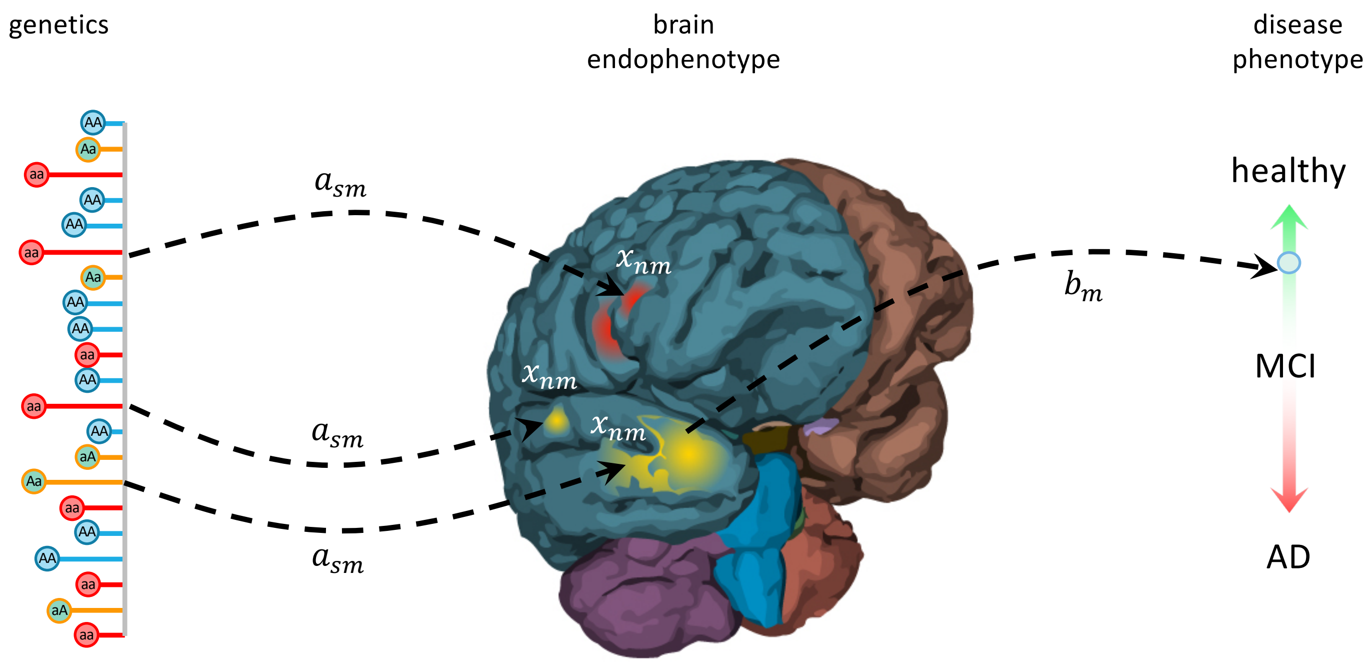
- Overview of Medical Imaging
  - Utility and properties
- Example: Segmentation
  - *Classical* and deep learning approaches
- Example: Registration (alignment):
  - Optimization and learning approaches
- Example: Imaging genetics
- Takeaways

Takeaway Goals

- Problems
  - Help the clinicians or scientists (don’t replace them)
- Tools and approaches
  - Probabilities, convolutions, and anatomical models
  - Clinical interpretation
- Challenges
  - The systems don’t really work (yet)
- Opportunity
  - Impact healthcare (and research)!

Medical Imaging

- Crucial tool in clinical practice
  - Diagnostic (and incidental findings)
  - Planning treatment
  - Guide small and large interventions
  - …



genetics
brain
endophenotype
disease
phenotype
AA
Aa
aa
aA
a_{sm}
a_{sm}
a_{sm}
x_{nm}
x_{nm}
x_{nm}
b_m
healthy
MCI
AD

Medical Imaging

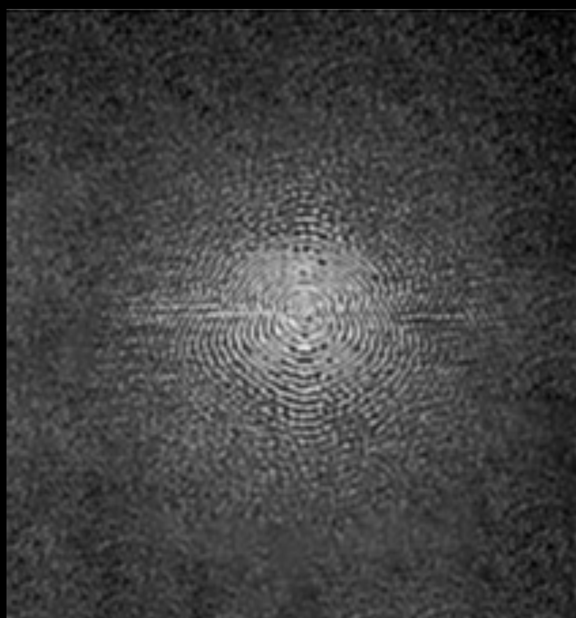
- Crucial tool in clinical practice
  - Diagnostic (and incidental findings)
  - Planning treatment
  - Guide small and large interventions
  - …
- Research
  - Clinical studies
  - Scientific studies

Medical Image **Analysis** (or: how can we help?)

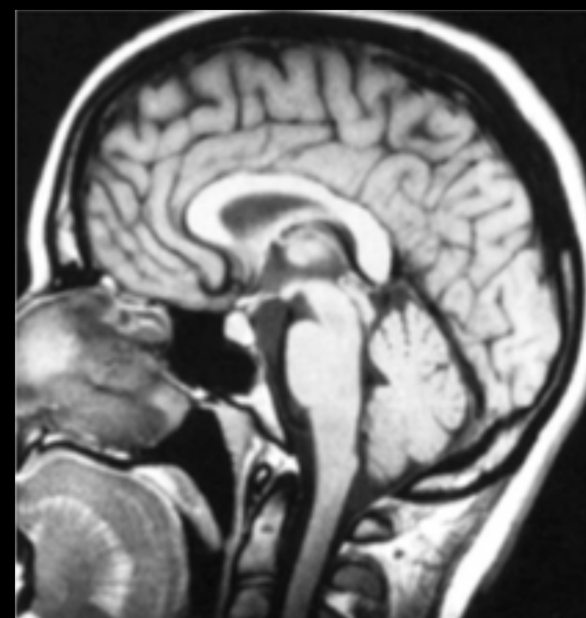
- Diagnosis algorithms - require large datasets
- Visualization - learn what to show, widely overlooked?
- **Segmentation** - outline, measure anatomy and pathology
- **Registration** - alignment for treatment planning, population analysis
- Acquisition - faster, better
- Abnormality detection - pathology
- Shape modelling
- Joint inference with other clinical data
- …

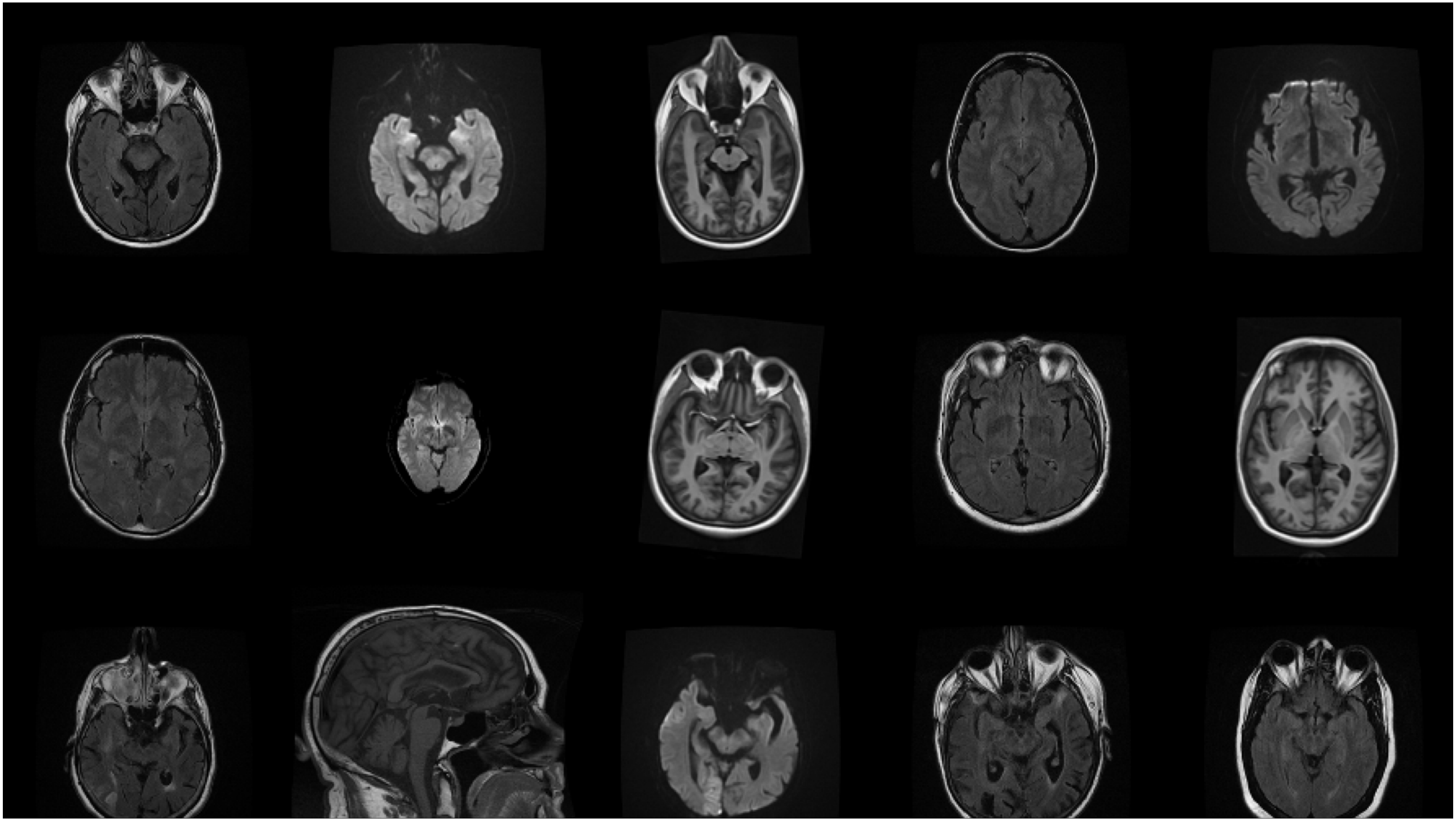
Properties of Medical Images

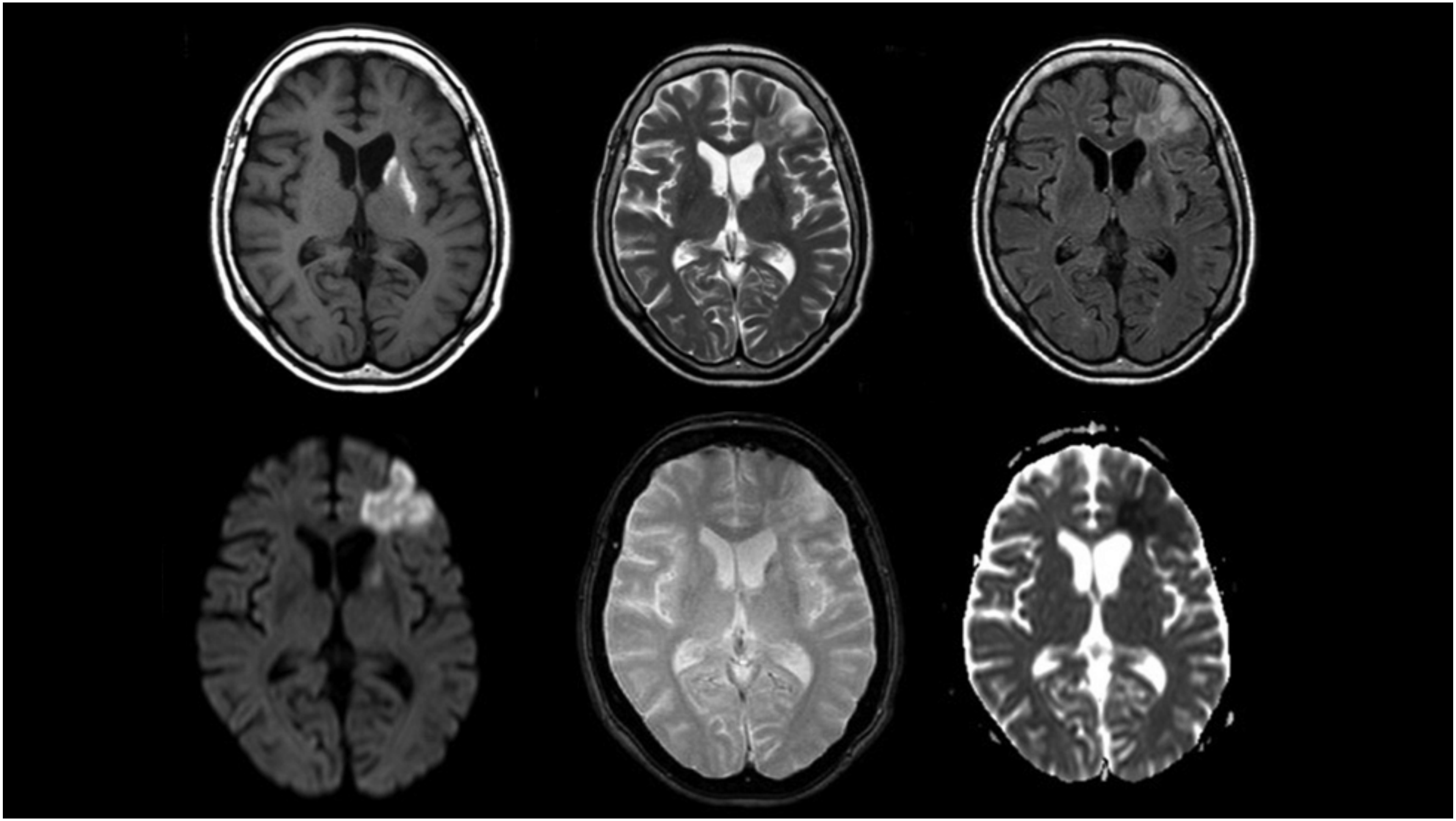
- Varies dramatically by image type

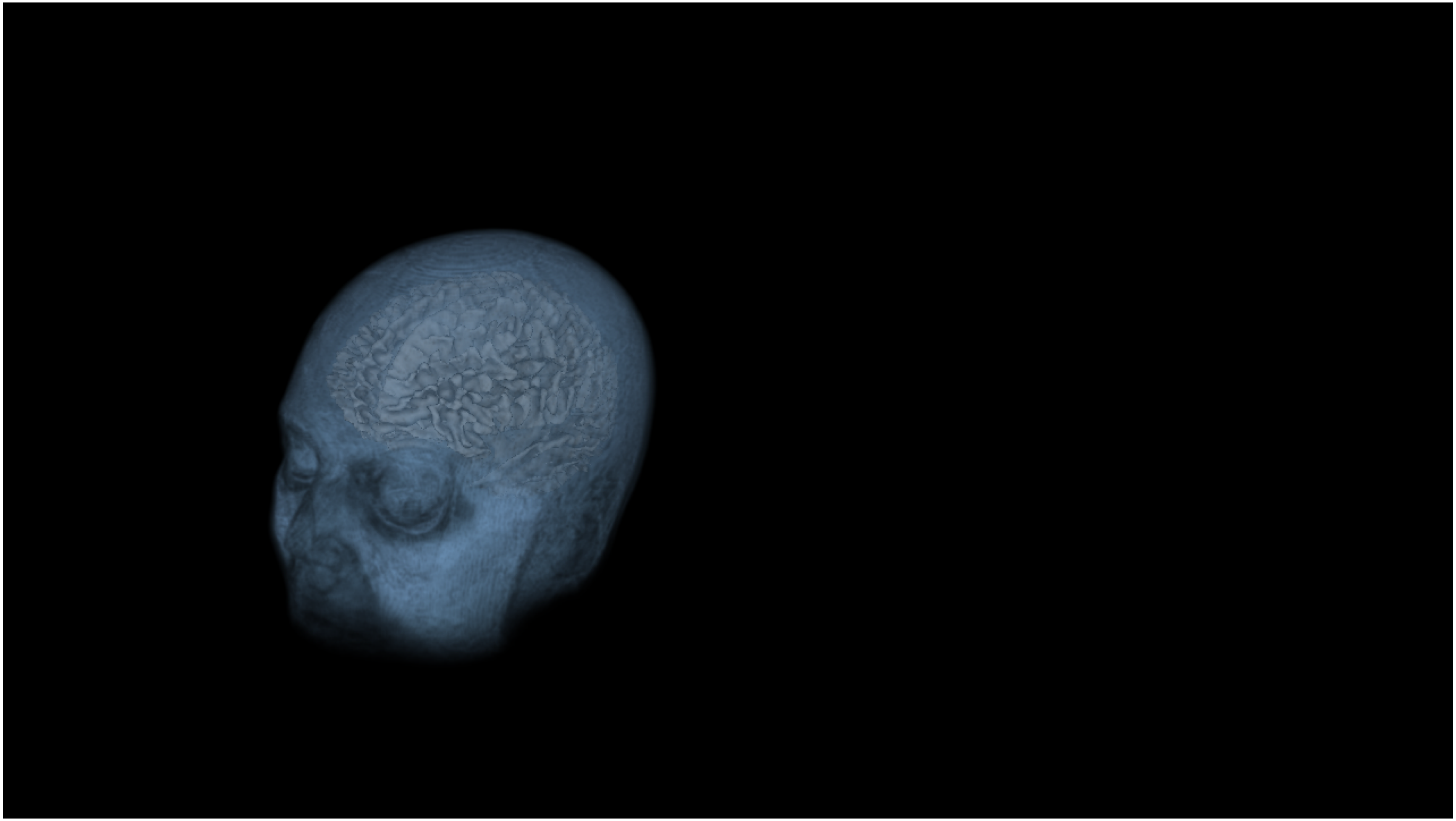


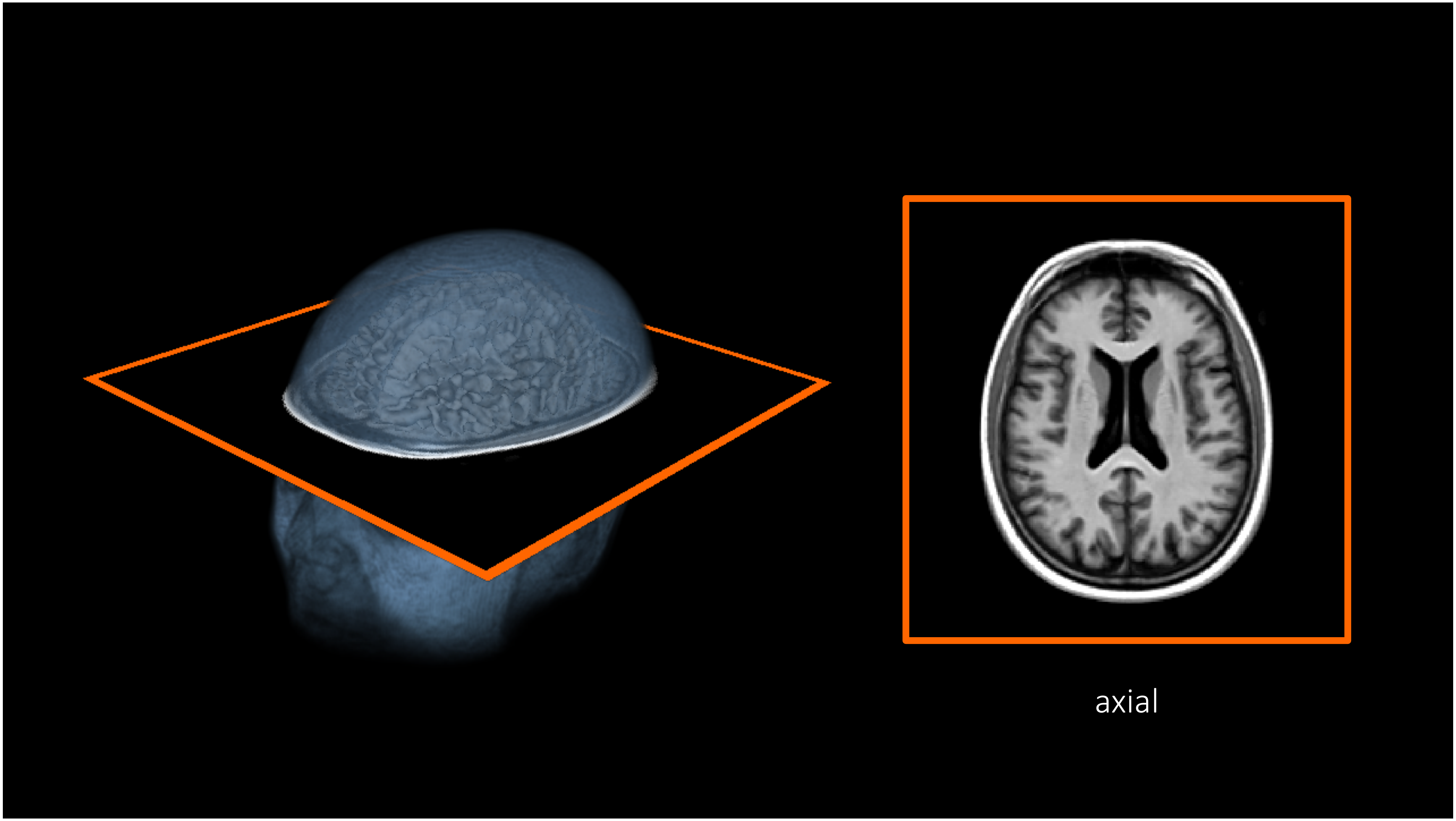
Fourier
Transform

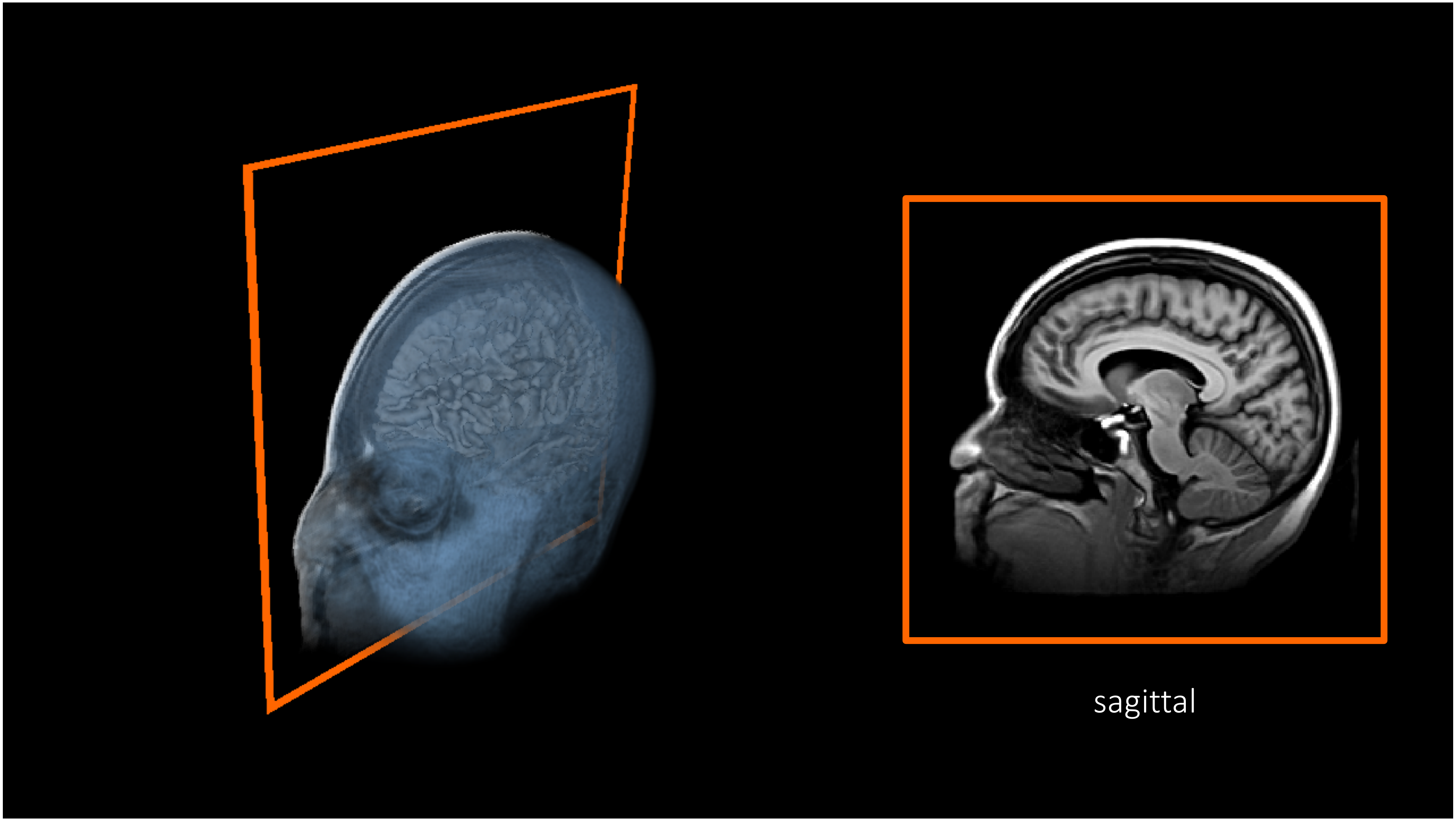


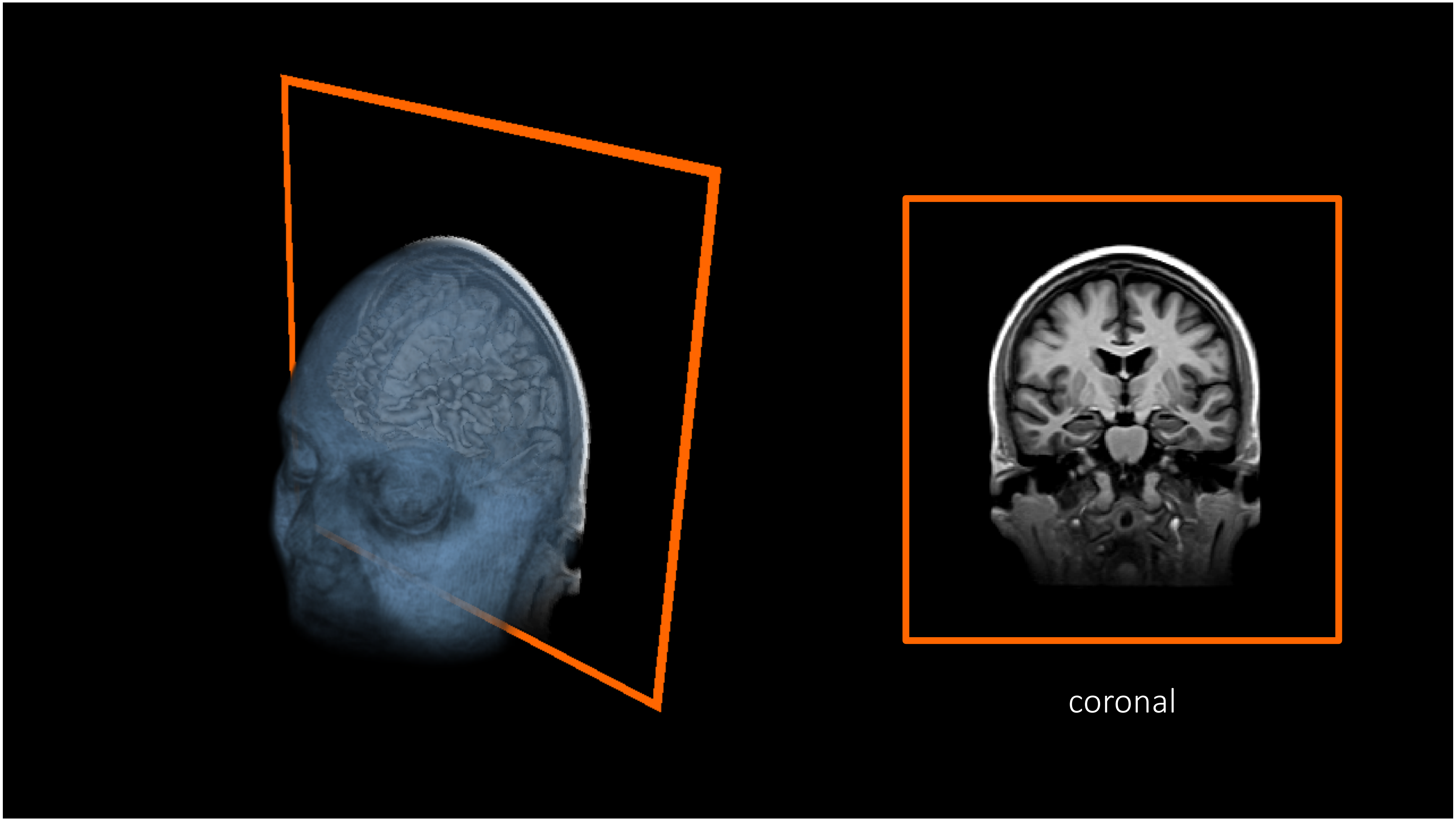




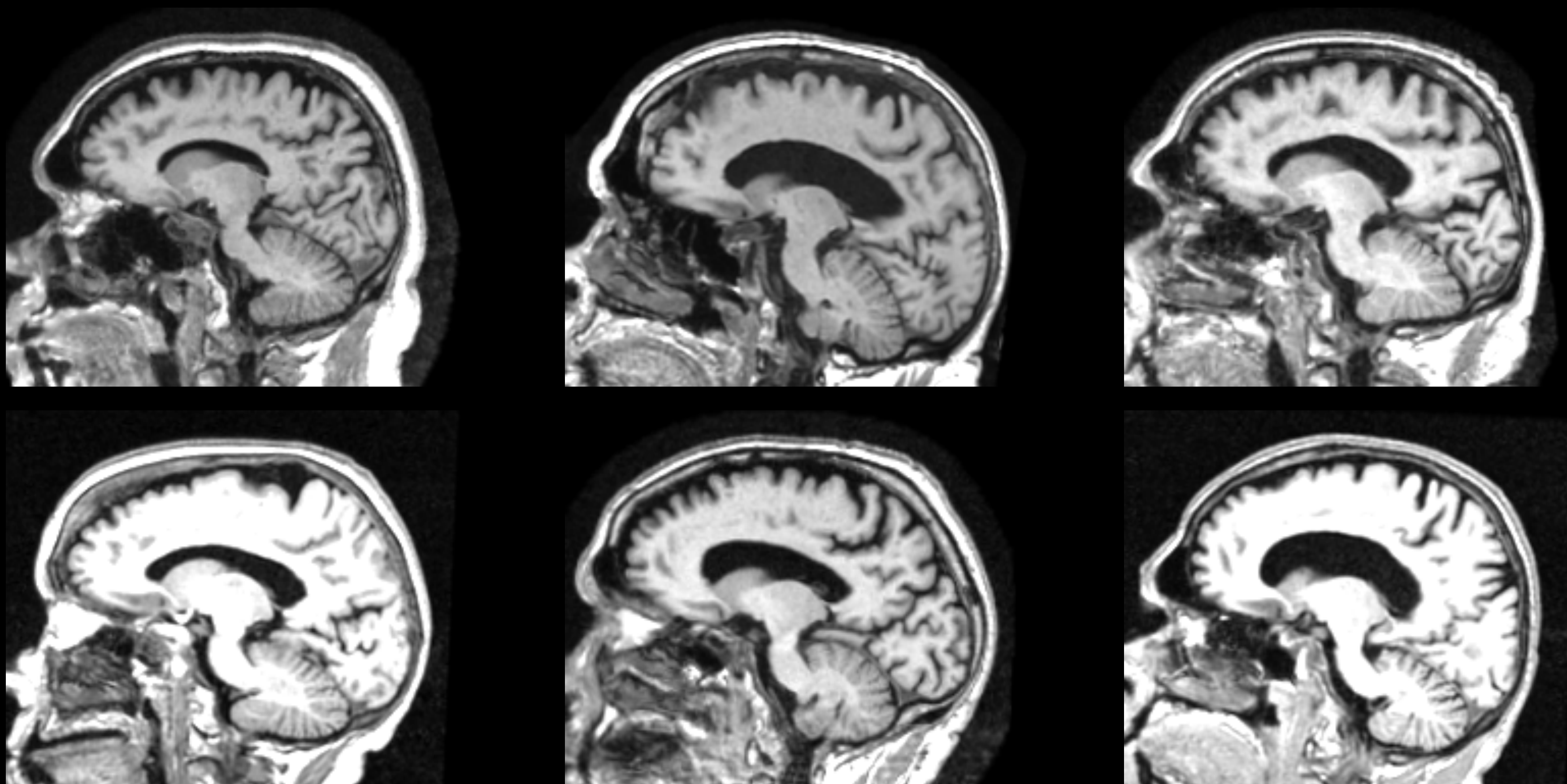



axial


sagittal


coronal

Variability and similarity



Properties

- Vary dramatically by image type
- MR Image quality:
  - Different noise patterns, patient motion, disease, many modalities
- Commonality of anatomy
- Pathology
  - can be big and obvious (e.g. tumor)…
  - … or very small and subtle (e.g. neurodegeneration)
- A lot of 3+ dimensions
  - So ‘**voxel**’ (**volume element**) instead of ‘pixel’ (picture element)

Questions?

Outline

- Overview of Medical Imaging
  - Utility and properties
- **Example: Segmentation**
  - ***Classical* and deep learning approaches**
- Example: Registration (alignment):
  - Optimization and learning approaches
- Example: Imaging Genetics
- Takeaways

Image Segmentation

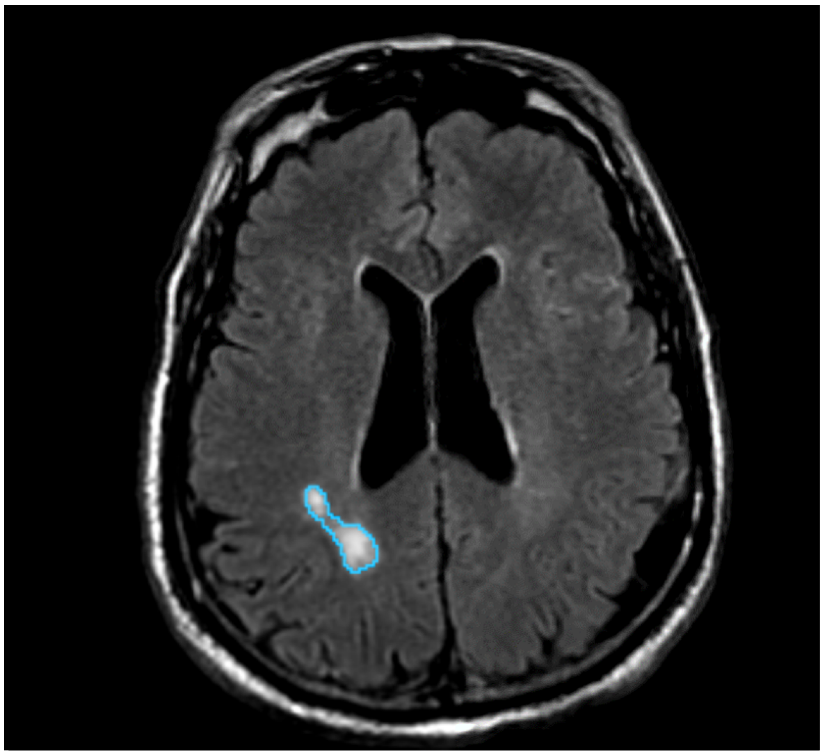
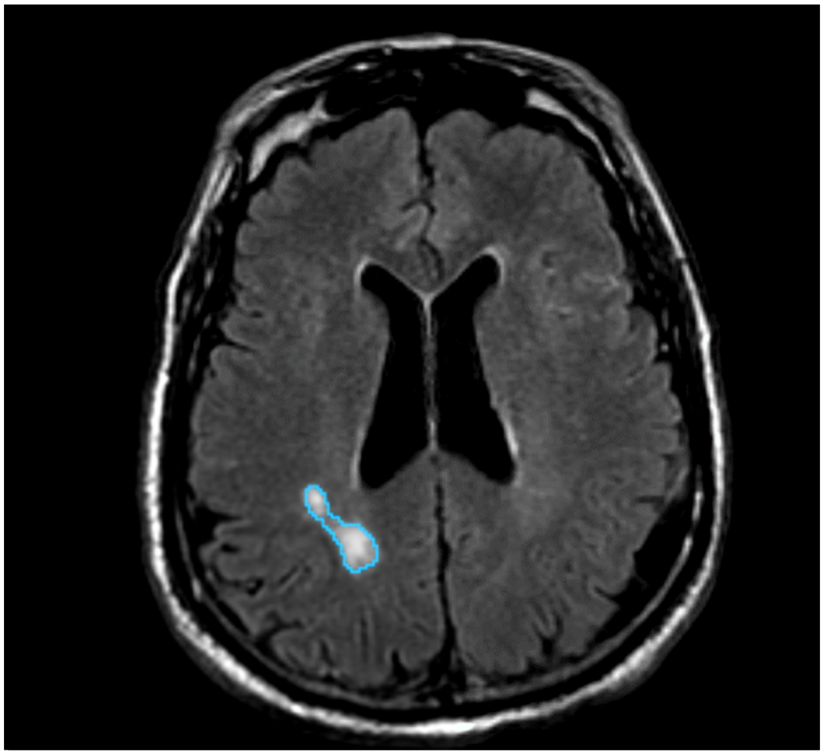
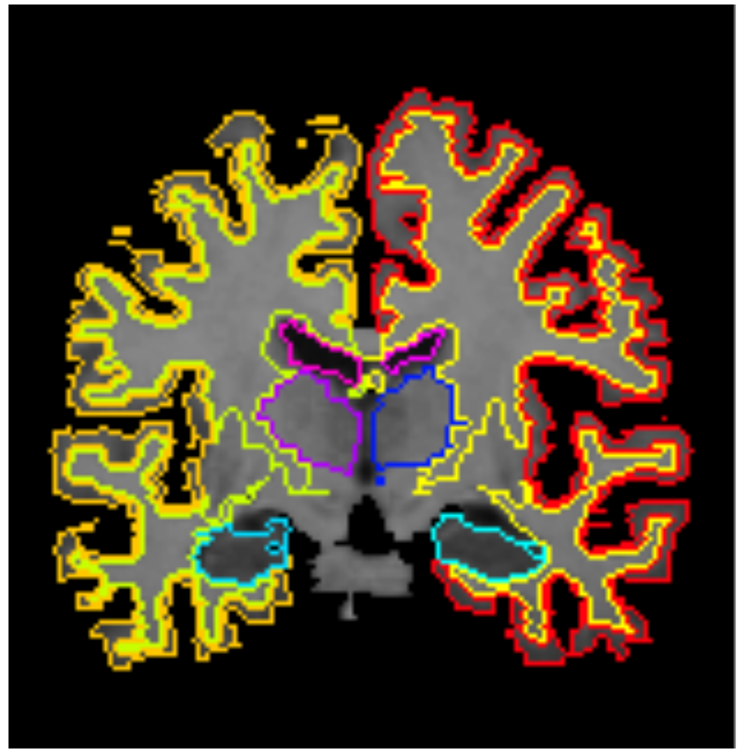
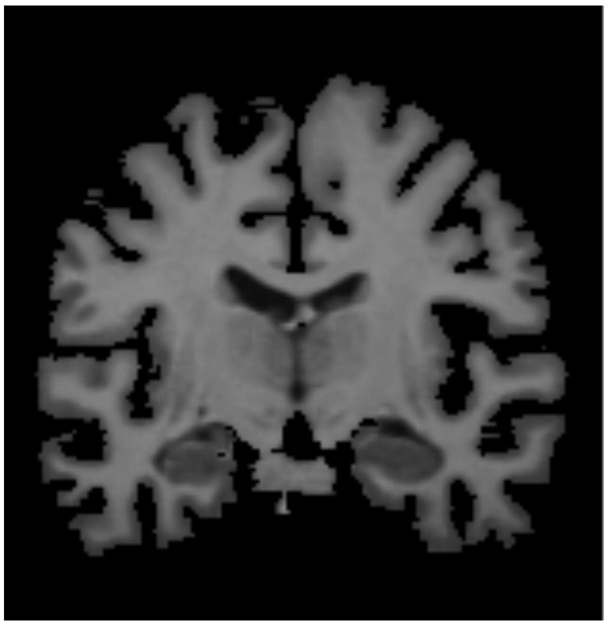


Image Segmentation

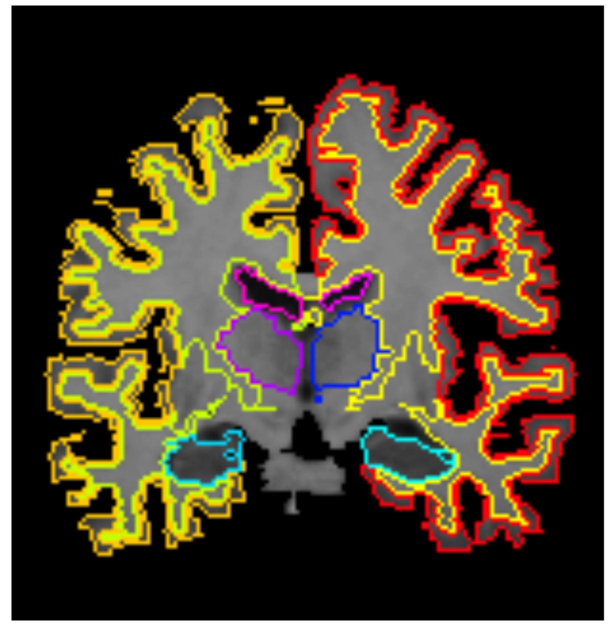




Supervised segmentation

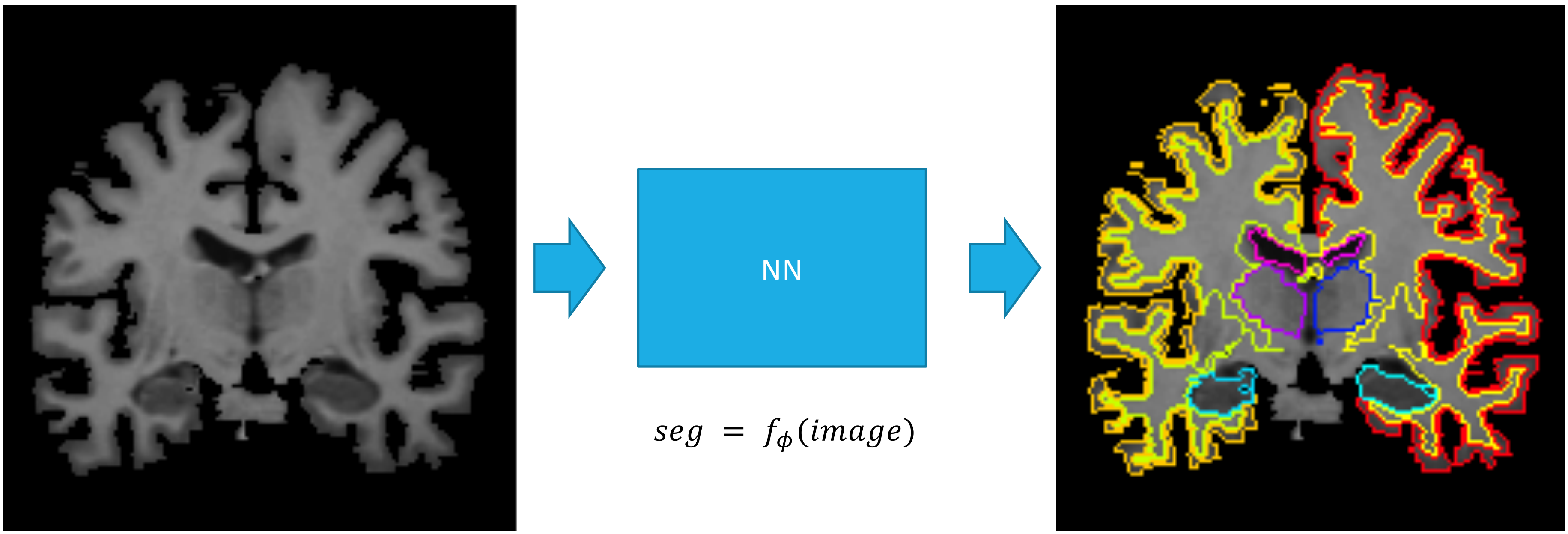


$$seg = f_{\phi}(image)$$



Supervised segmentation

Large example dataset: solved problem by DL?



$$seg = f_{\phi}(image)$$

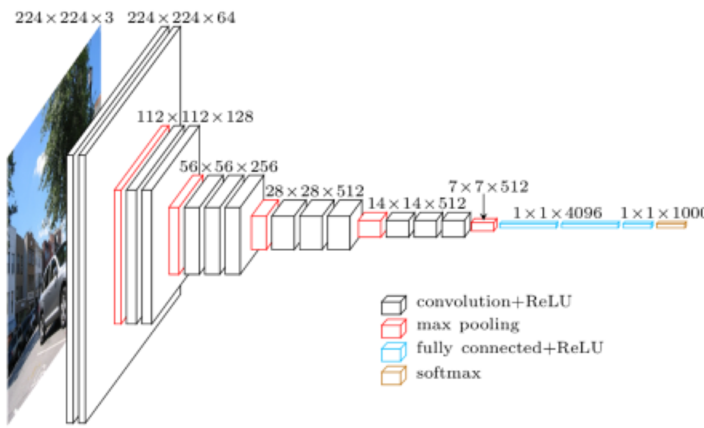
What kind of NN?

VGG, etc?

Use existing multi-label networks

Architecture: convolutions, max-pools, fully connected, etc.

But need to output 8 million voxels! – hours!



anatomical
label
(one-hot encoding)

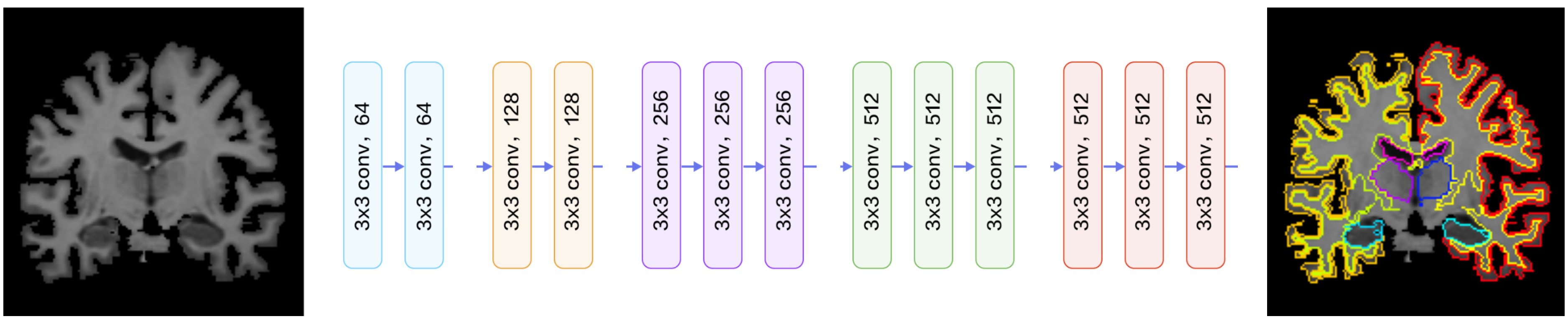
Fully convolutional?

Input-output both high dimensional, no max-pooling (make 3D)

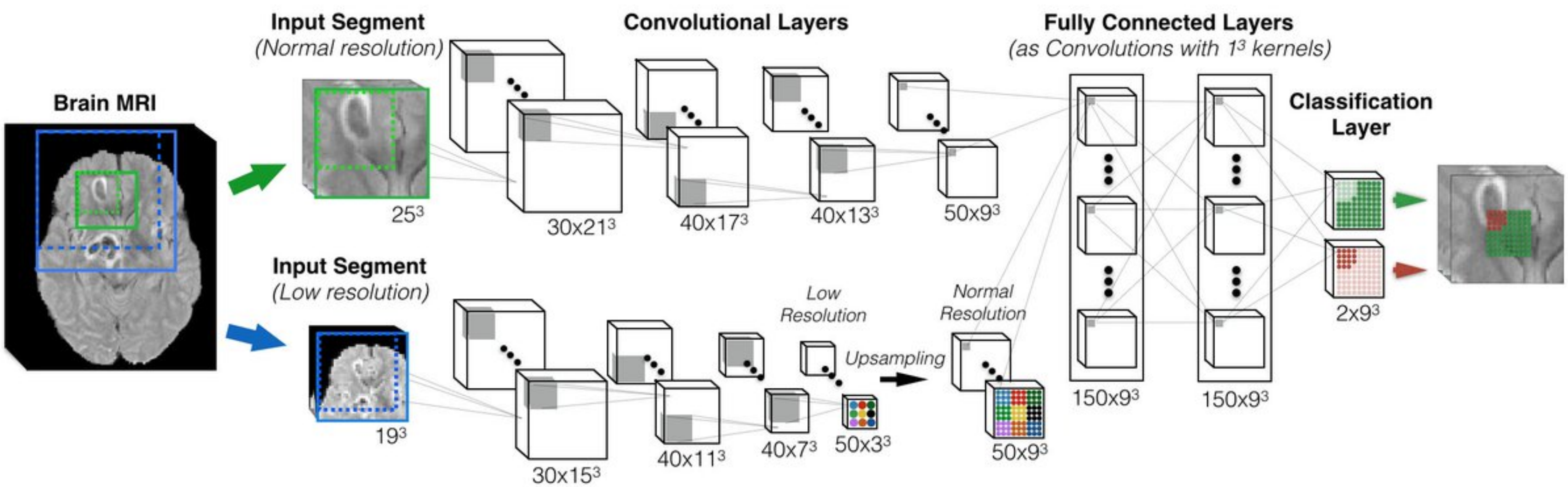
10-layer network: don’t have enough context to predict anatomy

Deep networks (100 layers) – require too many parameters

channels = # labels

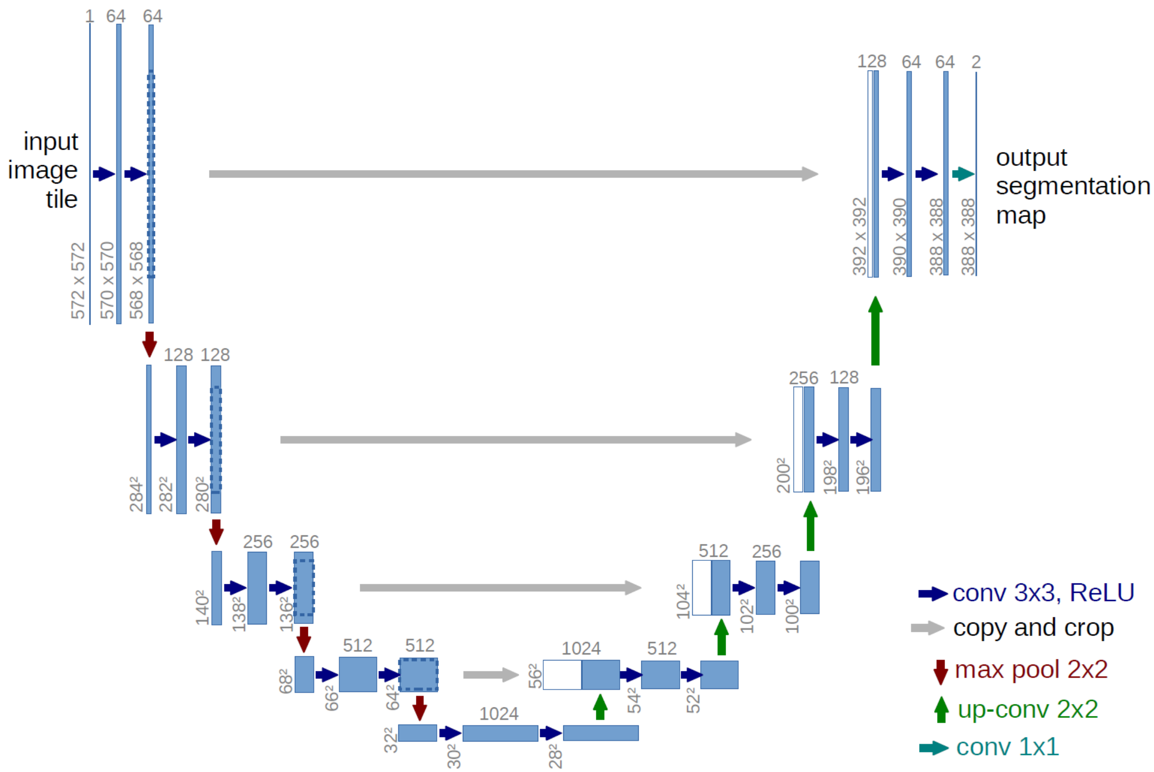


Multi-scale inputs



Kamnitsas, et al., MedIA 2016

U-Net



Ronneberger et al, 2015

What kind of CNN?

Network architecture

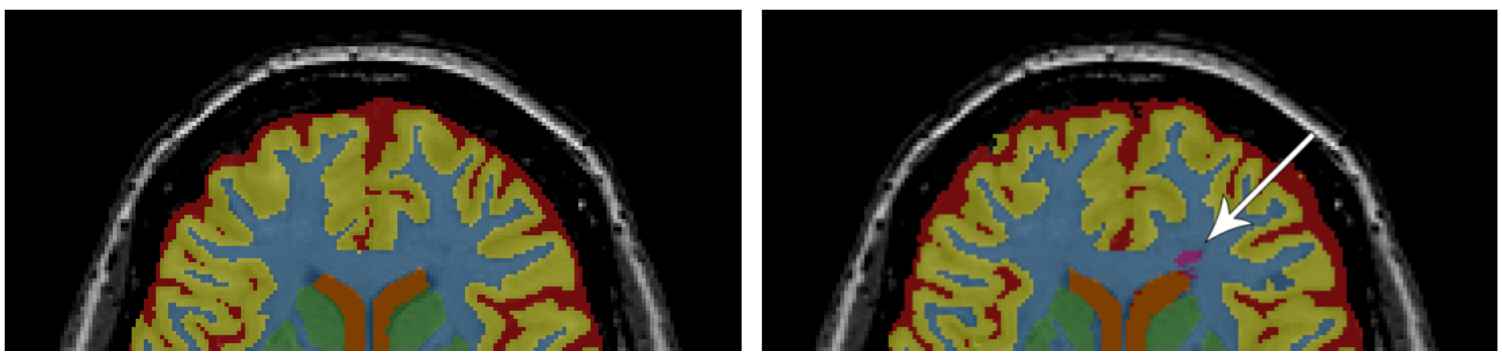
- Predict each voxel (e.g. 3D **VGG**)? too slow, cumbersome
- Fully Convolutional? Large memory, parameter space, not enough **field of view**
- Multiscale input?
- UNet!

Results

Dice (Volume Overlap)	Dice (Volume Overlap)	Runtime
FreeSurfer (classical state of the art)	~80	~6-24 hours
Deep Methods	~85-91	~1 second-1 hour

Problems

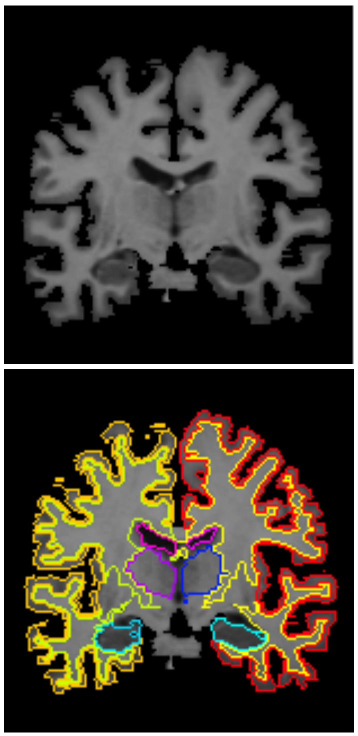
- Often don’t actually have these **segmented** data
  - Long time to segment for experts!
  - Too many modalities
  - Too much variation (especially pathologies)
- Our metrics
  - Easy to compute, differentiate
  - Often not **anatomically** meaningful

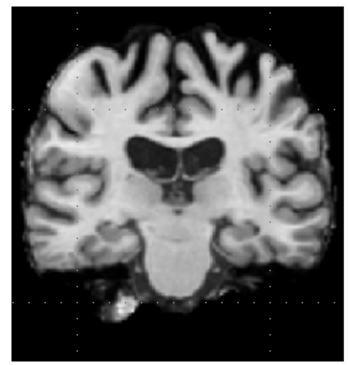


P. Moeskops et al, DLMIA, 2017

Segmentation in a more realistic setting

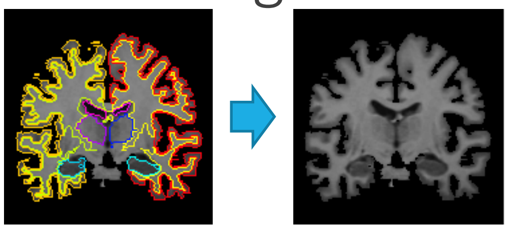
Few (one) segmented example

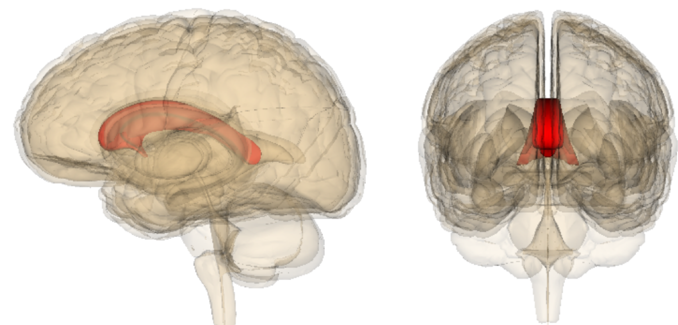




Probabilistic (Generative) Model

- Define segmentation -> image model $p(I|S) * P(S)$



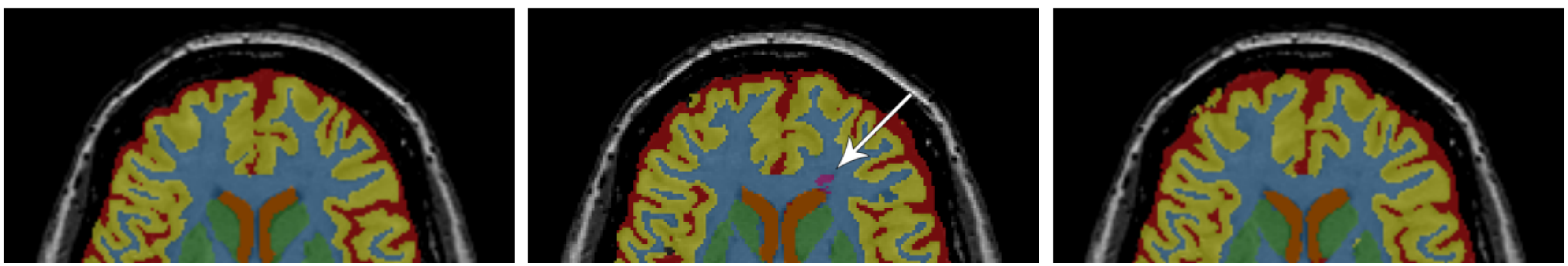


- Enables knowledge (priors) into segmentation model $p(S)$
  - $p(S)$ defined based on likely *shapes* of each label
  - $P(I|S)$ is the intensity (distribution) for each label
  - **Inference**: $p(S|I)$ at each voxel: **label** matches the **intensity** such that **shapes** make sense.

Probabilistic (Generative) Model

- Combine with deep learning predictions:

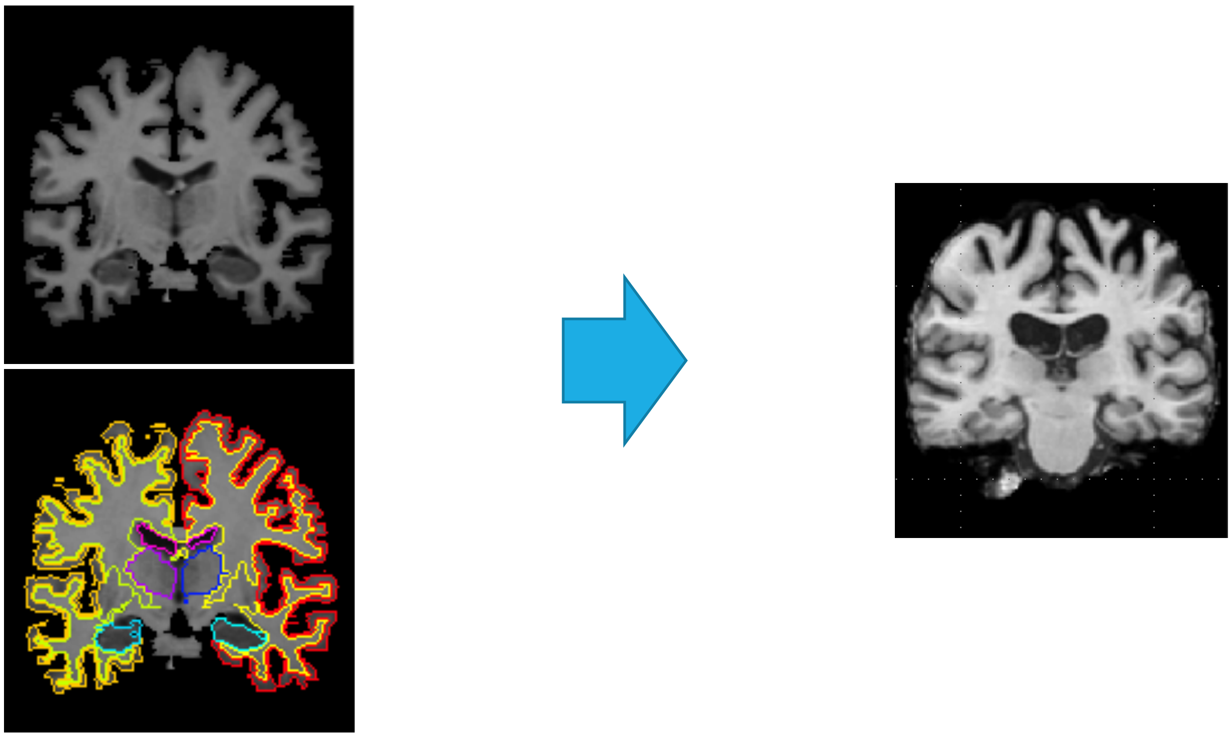
  p(S) can be anatomically specified
  or learned from another distribution
- Attach prior to network, or modelling through VAEs, etc…



P. Moeskops et al, DLMIA, 2017

Brains are similar!

- Can similarity of brains help?



Questions?

Outline

- Overview of Medical Imaging
  - Utility and properties
- Example: Segmentation
  - *Classical* and deep learning approaches
- **Example: Registration (alignment):**
  - **Optimization and learning approaches**
- Example: Imaging Genetics
- Takeaways

Image Registration

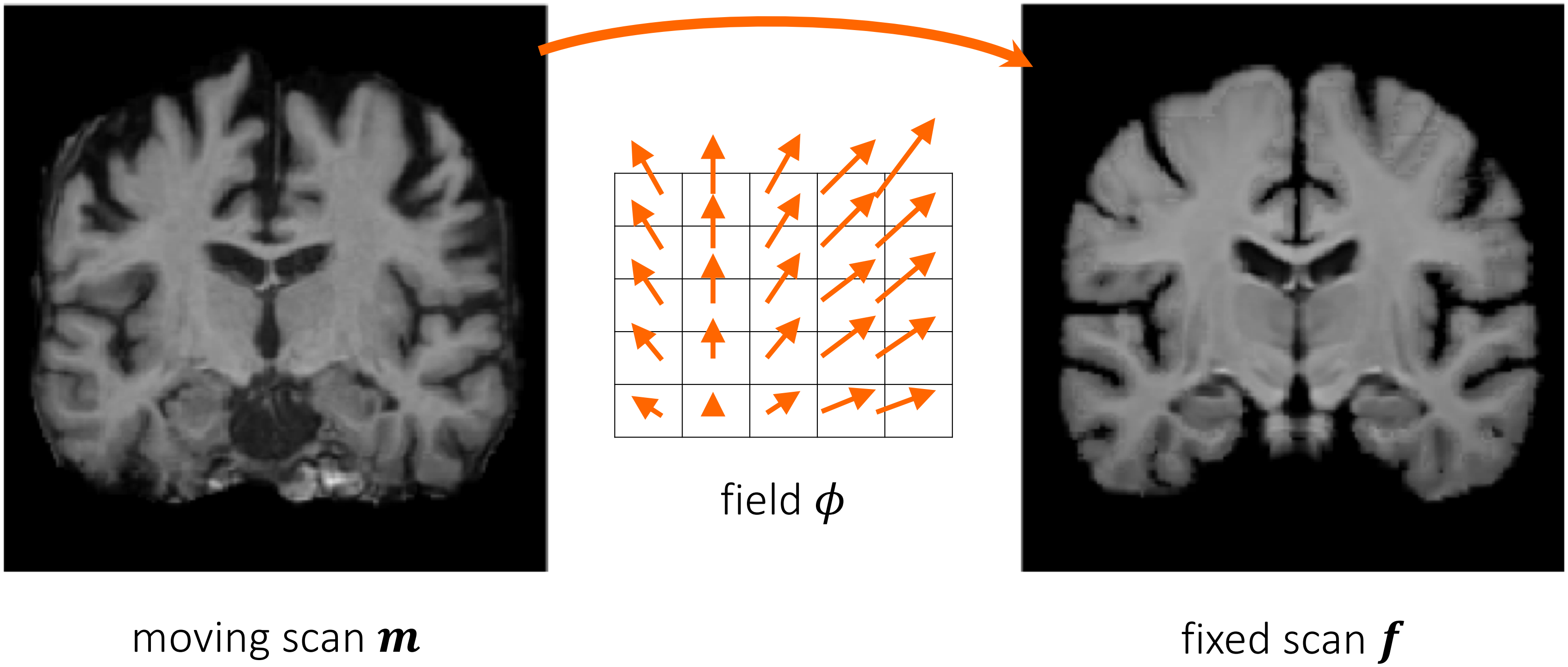
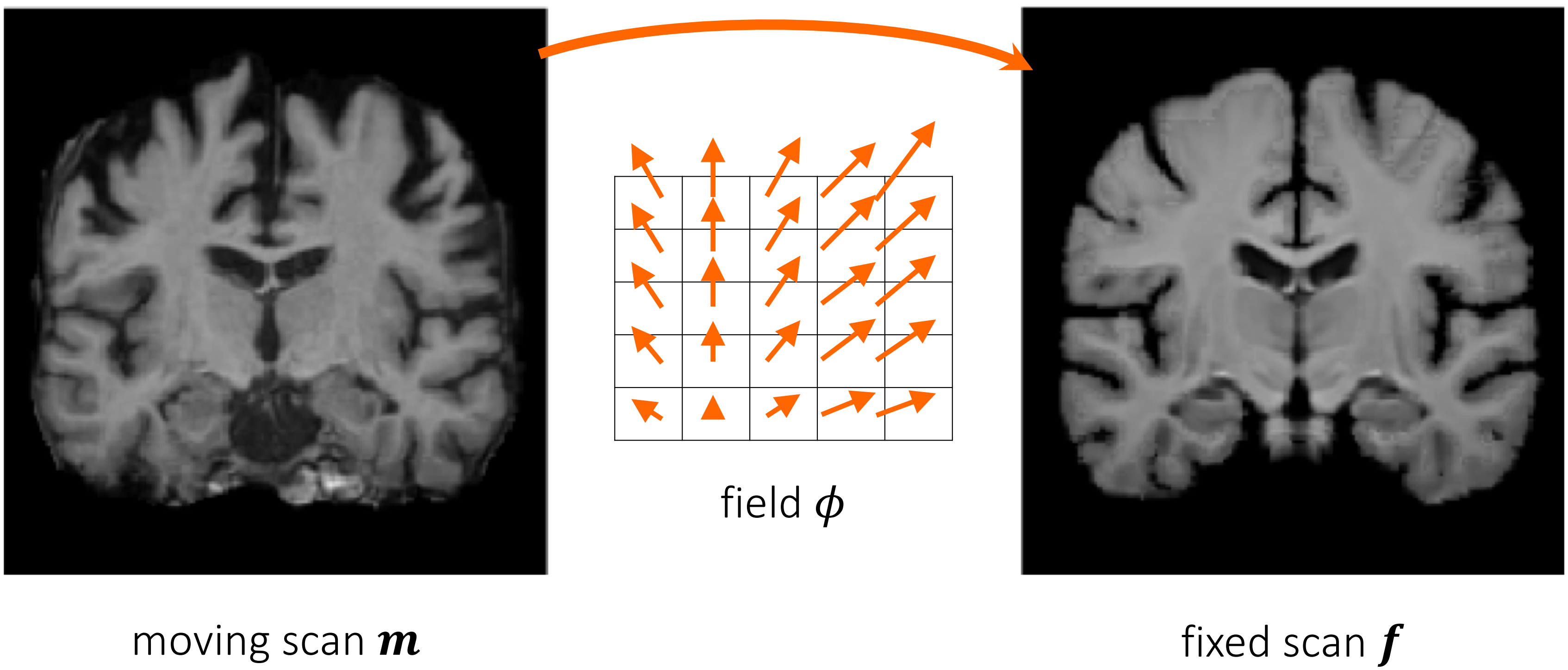
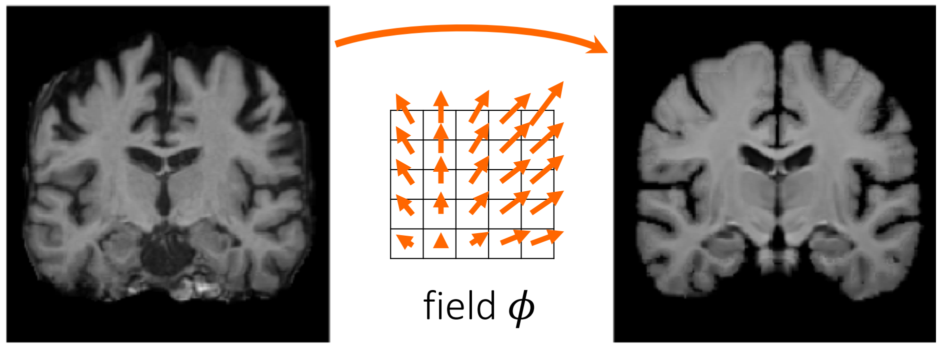


Image Registration



Traditional approach

scan $\boldsymbol{m}$



scan $\boldsymbol{f}$

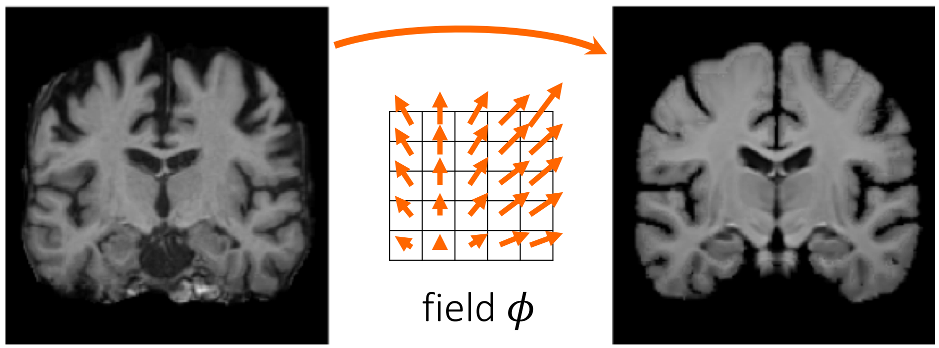
warped scan m

$$\hat{\phi}_{m,f} = argmin_{\phi} \underbrace{\|m \circ \phi - f\|}_{\text{images match}}$$

optimal deformation field

Traditional approach

scan $\boldsymbol{m}$



scan $\boldsymbol{f}$

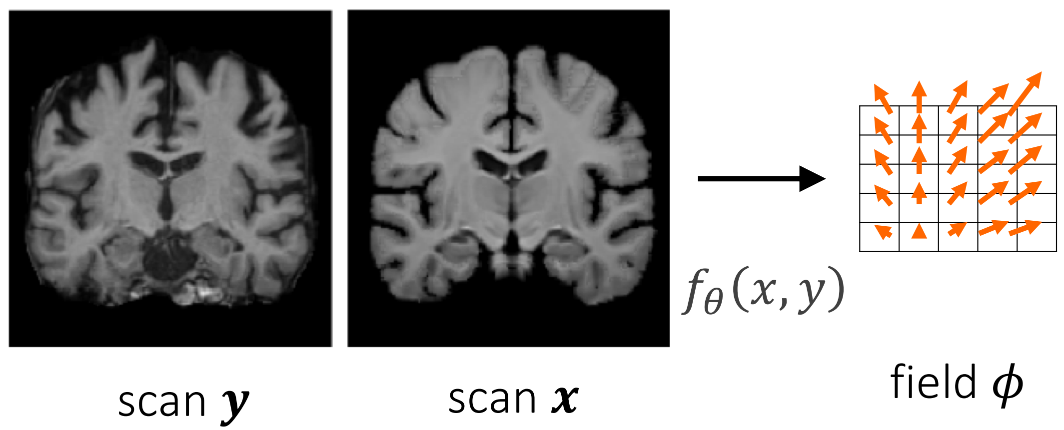
warped scan m

$$\hat{\phi}_{m,f} = argmin_{\phi} \underbrace{\|m \circ \phi - f\|}_{\text{images match}} + \lambda \underbrace{\|\nabla\phi\|}_{\text{Anatomically smooth}}$$

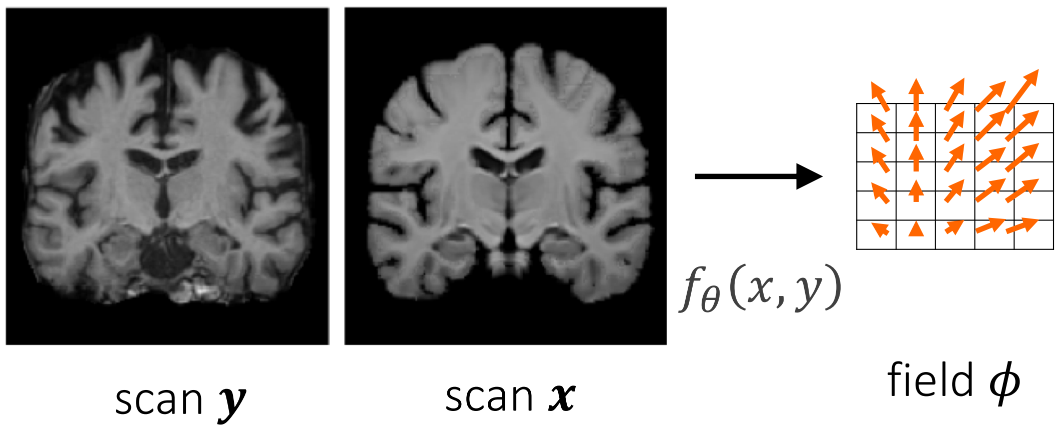
optimal deformation field

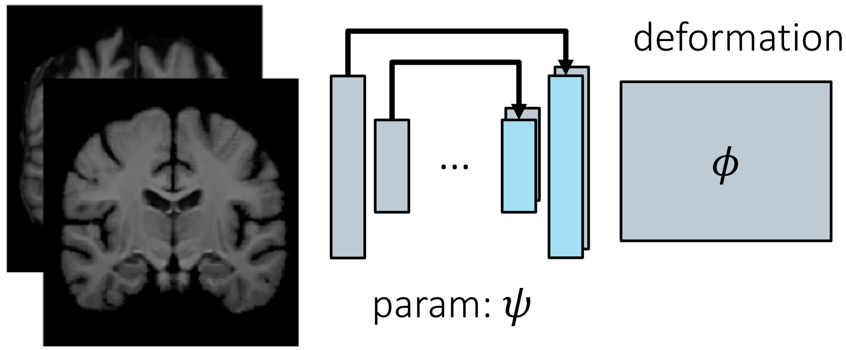
Pairwise optimization: slow (hours per image on CPU)

How can machine learning help?

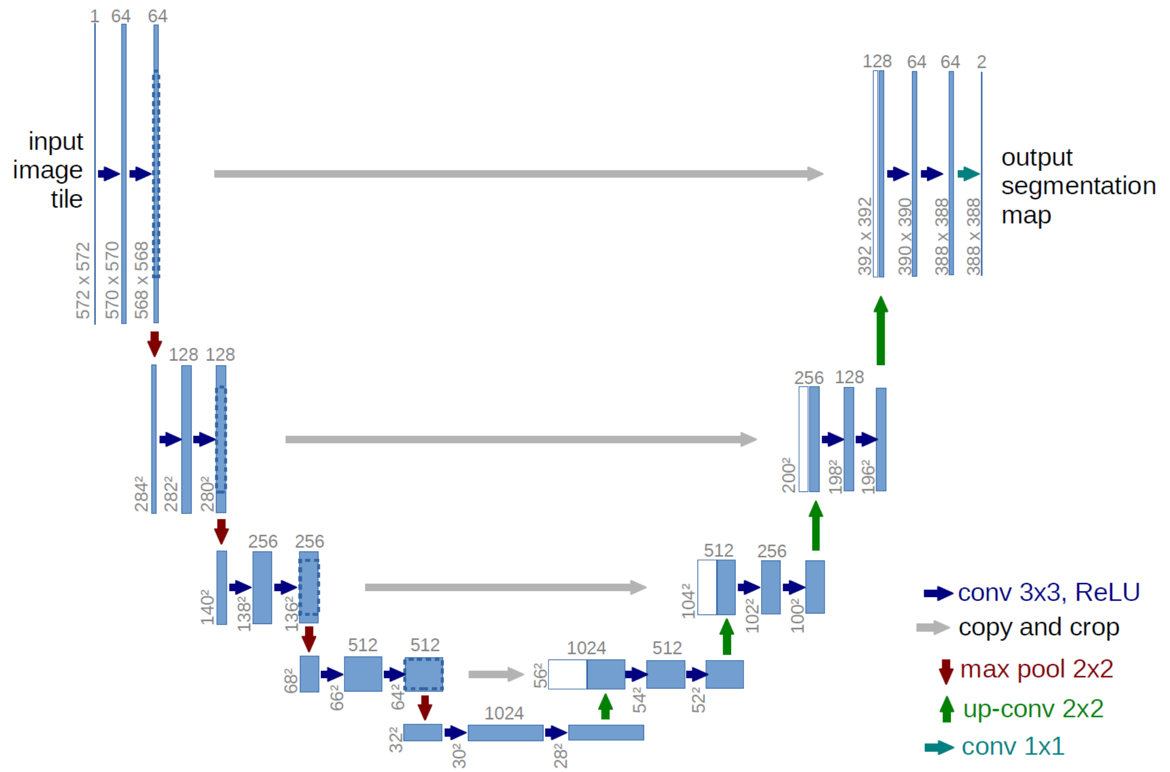


Supervised Learning



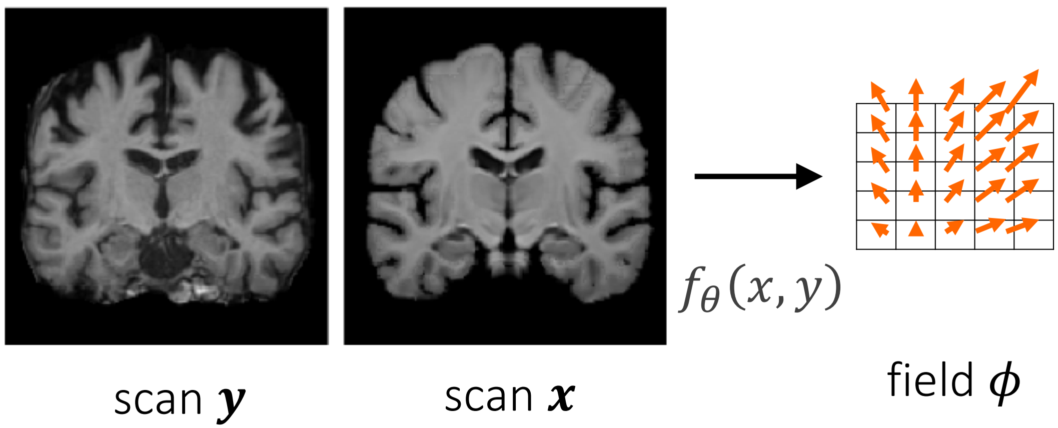


What kind of architecture?



Ronneberger et al, 2015

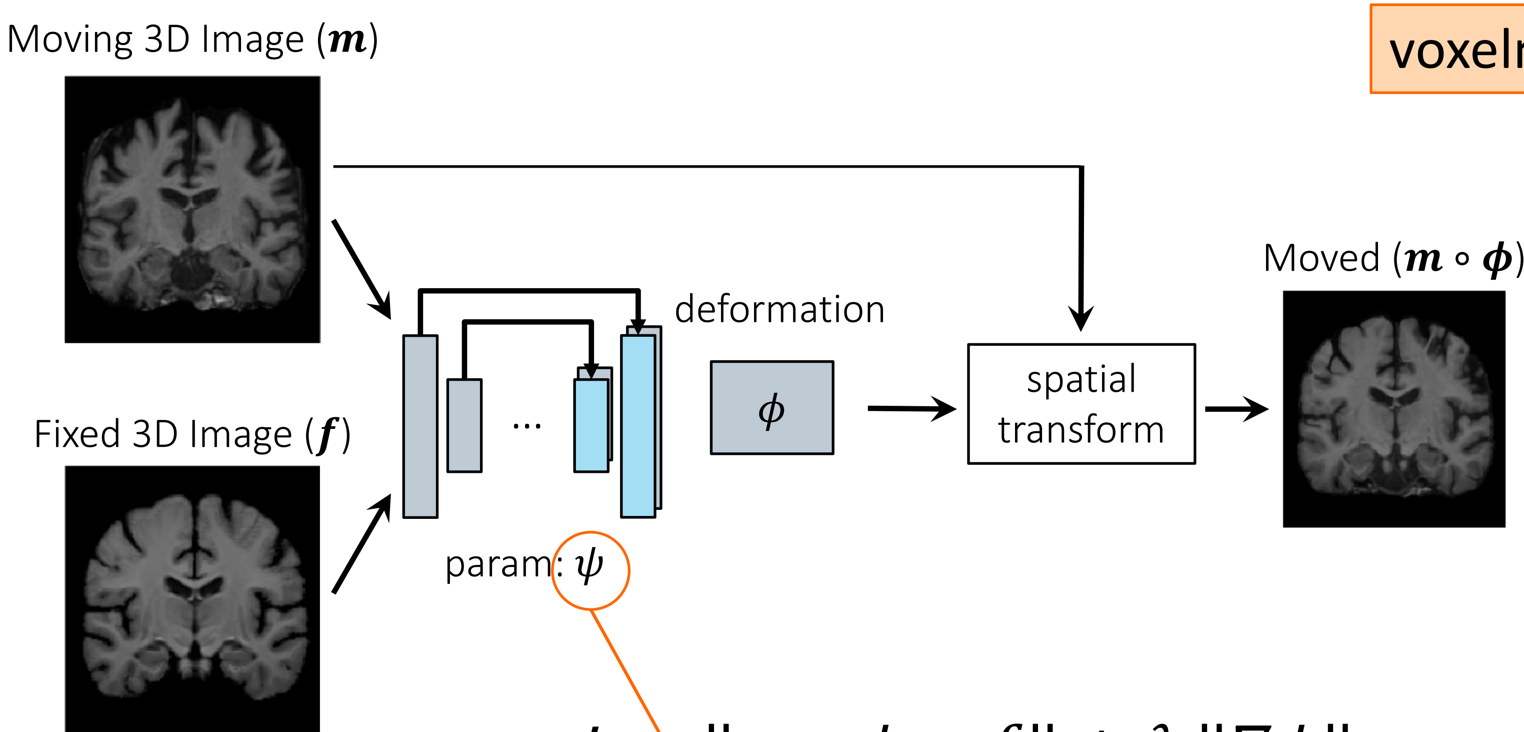
Supervised Learning



fast for new image pair!
need **ground truth** registration $\boldsymbol{\phi}$

Unsupervised Learning: VoxelMorph

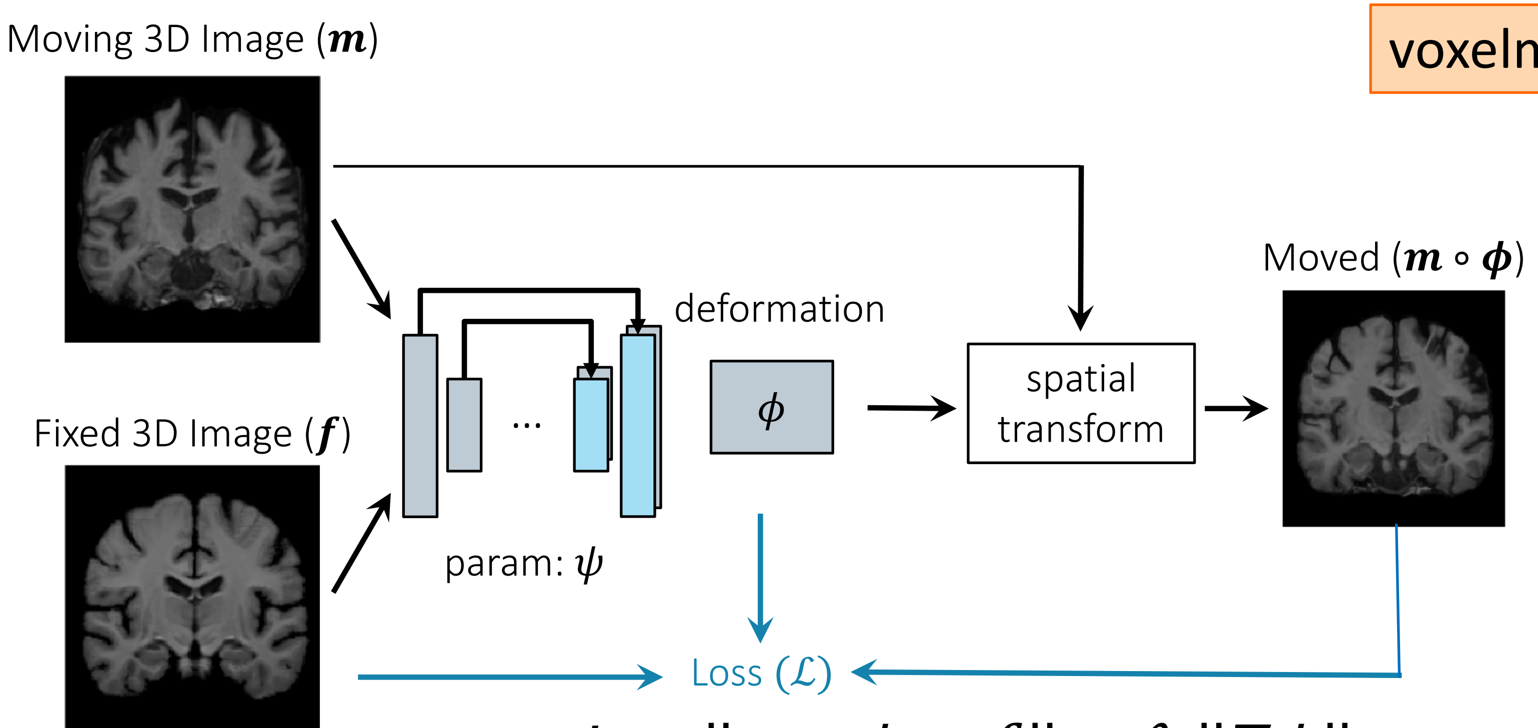
voxelmorph.mit.edu



$$\min_{\psi} \underbrace{\|m \circ \phi - f\|}_{\text{images match}} + \lambda \underbrace{\|\nabla \phi\|}_{\text{Anatomically smooth}}$$

Unsupervised Learning: VoxelMorph

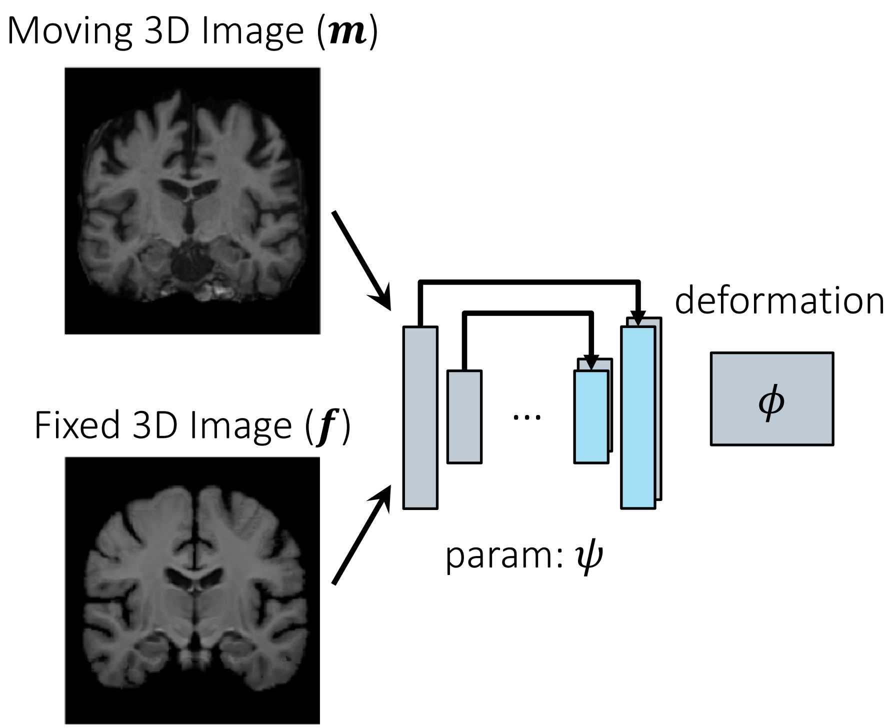
voxelmorph.mit.edu



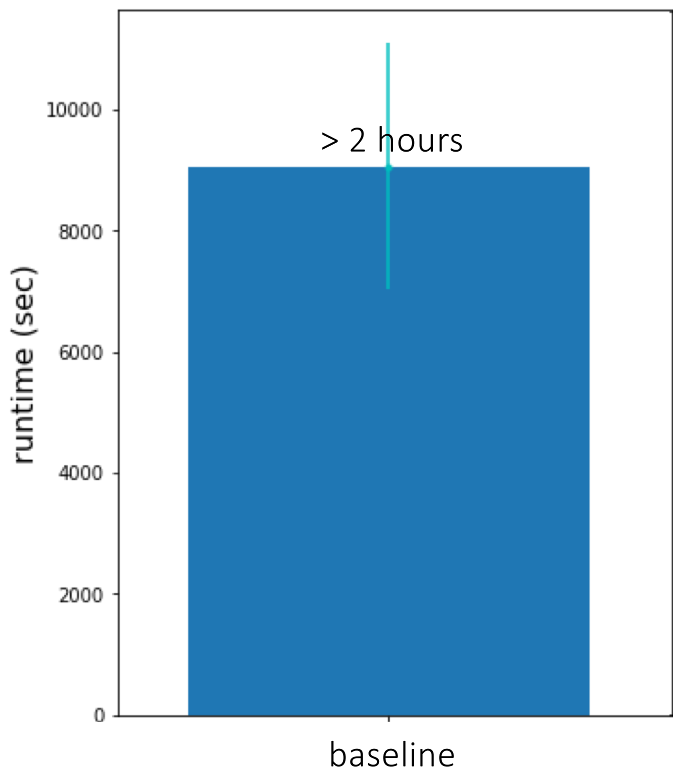
$$min_{\psi} \underbrace{||m \circ \phi - f||}_{\text{images match}} + \lambda \underbrace{||\nabla \phi||}_{\text{Anatomically smooth}}$$

Balakrishnan et al
CVPR 2018, TMI 2019

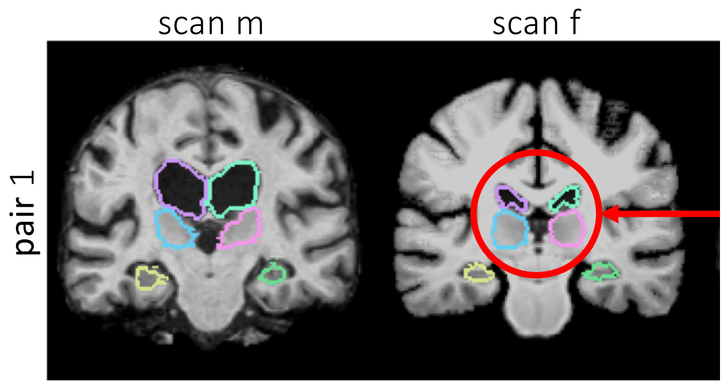
Registering a new image pair



Runtime for a new 3D image pair

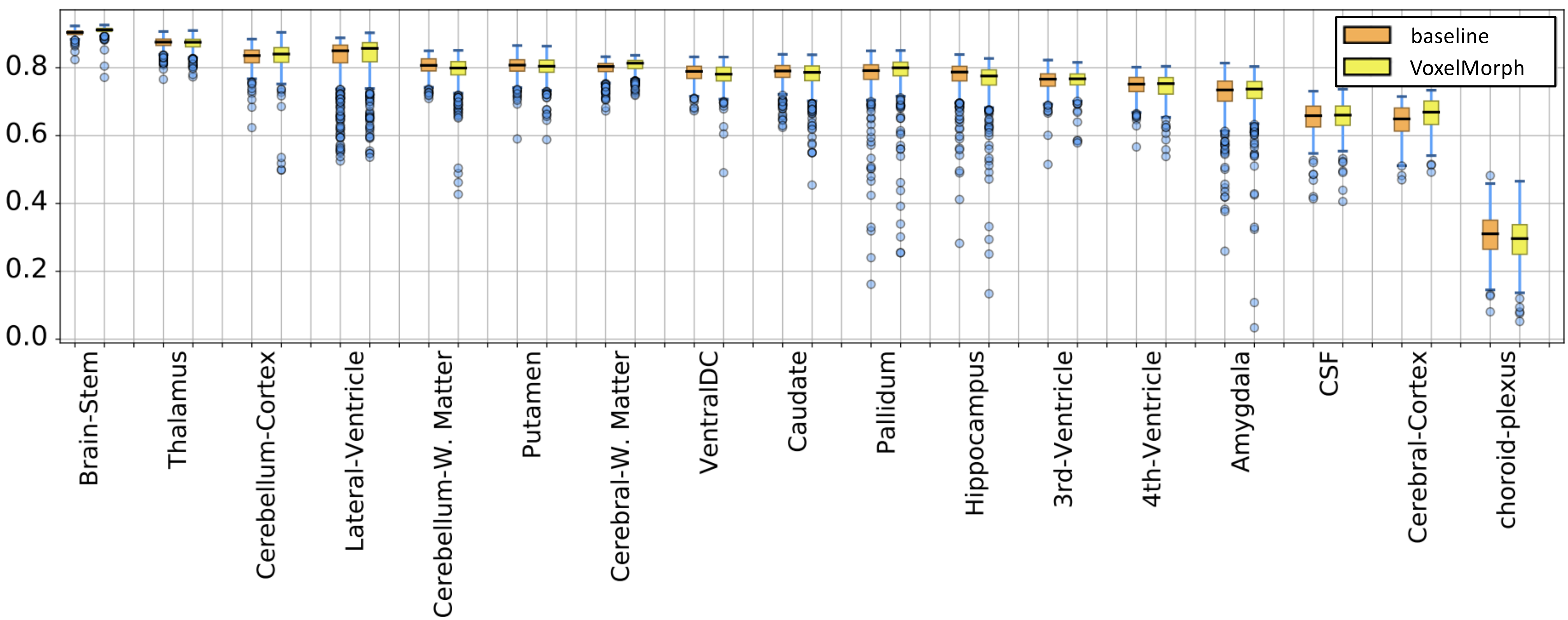


How to evaluate?



*algorithms only see images, no segmentation maps

Accuracy via volume overlap (Dice)



Remarks

- We derive network probabilistically **from probabilistic model**
  - $p(m|\phi; f) * p(\phi) \rightarrow p(\phi|m; f)$
  - Variational approximation to $p(\phi|m; f)$ leads to network

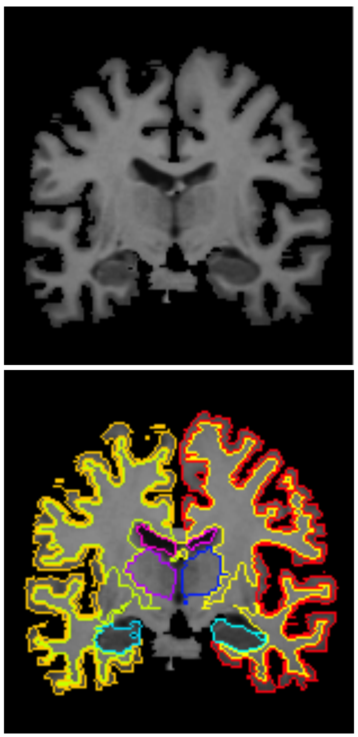
- Can impose stricter anatomical consistency (**diffeomorphisms**)
  - Provide topological guarantees

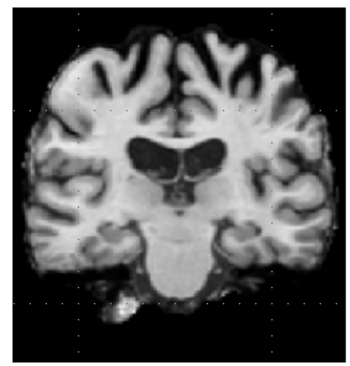
- Can use segmentations during training if we have them.

Dalca et al, MICCAI 2018

Going back to segmentation…

- Can similarity of brains help?





Questions?

Caveat: registration isn't perfect

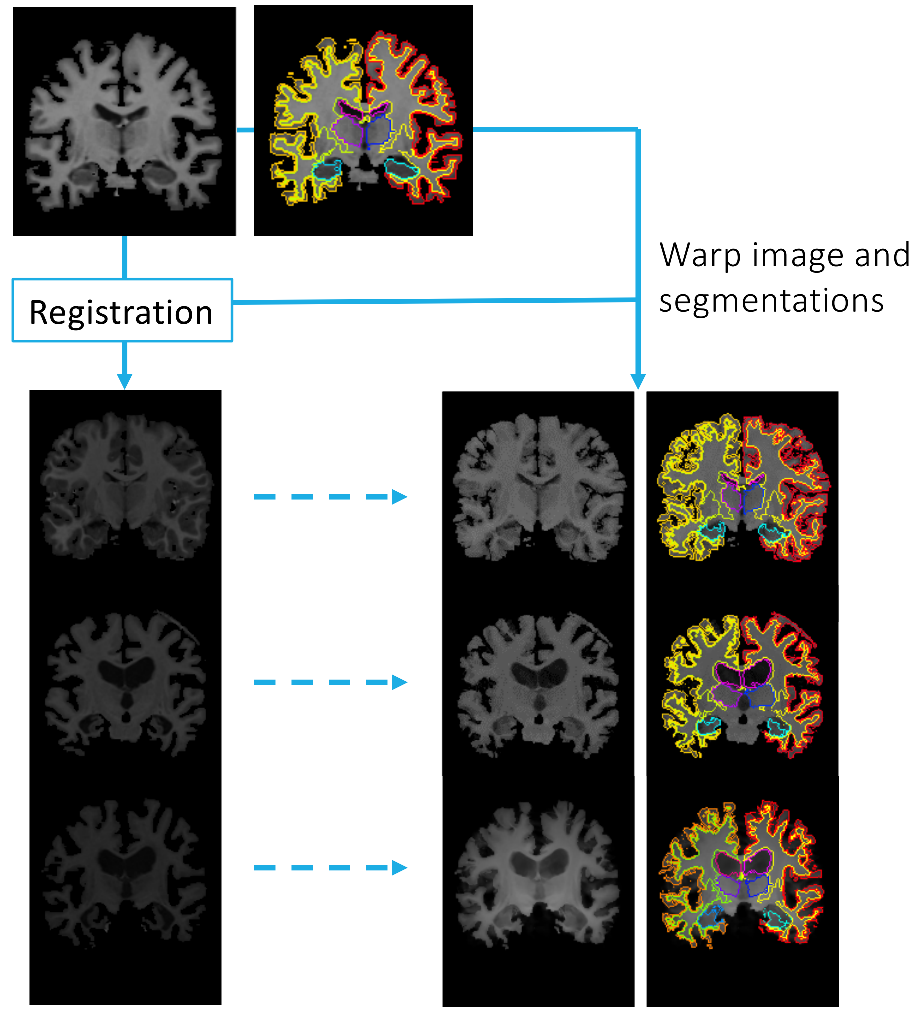
- **Supervised** segmentation (with 200 training images): 85 ± 9
- Registration-based segmentation (with 1 training image): 76 ± 14

Caveat: registration isn't perfect

- **Supervised** segmentation (with 200 training images): 85 ± 9
- Registration-based segmentation (with 1 training image): 76 ± 14
- Combine advantages!

Supervised segmentation & registration

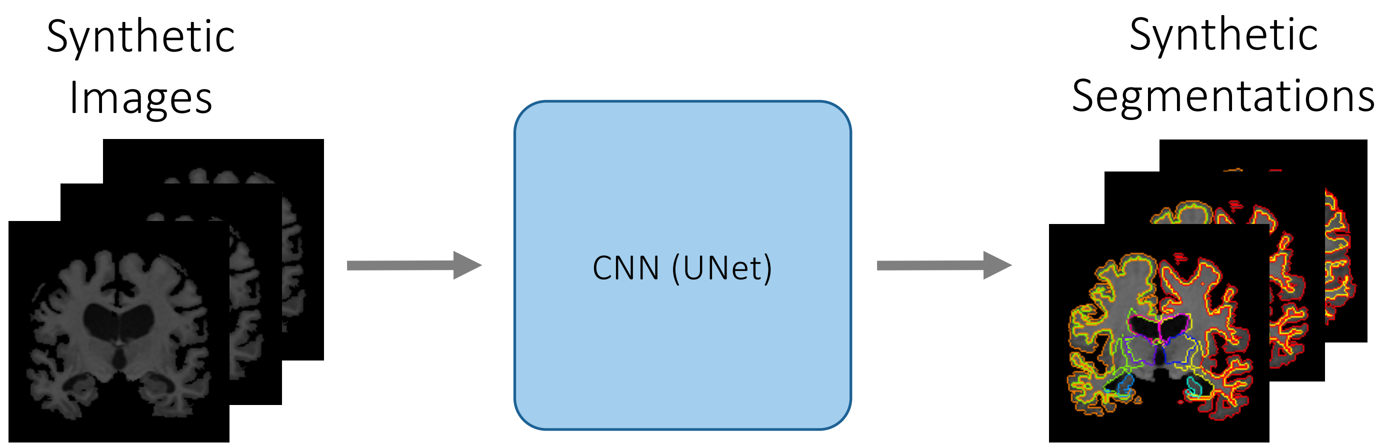
- Register training image to every image in dataset
  - distribution of transforms
- Warp labelled scan and segments to produce *supervised dataset*
  - Span anatomical distribution
  - Accurately segmented



CVPR2019

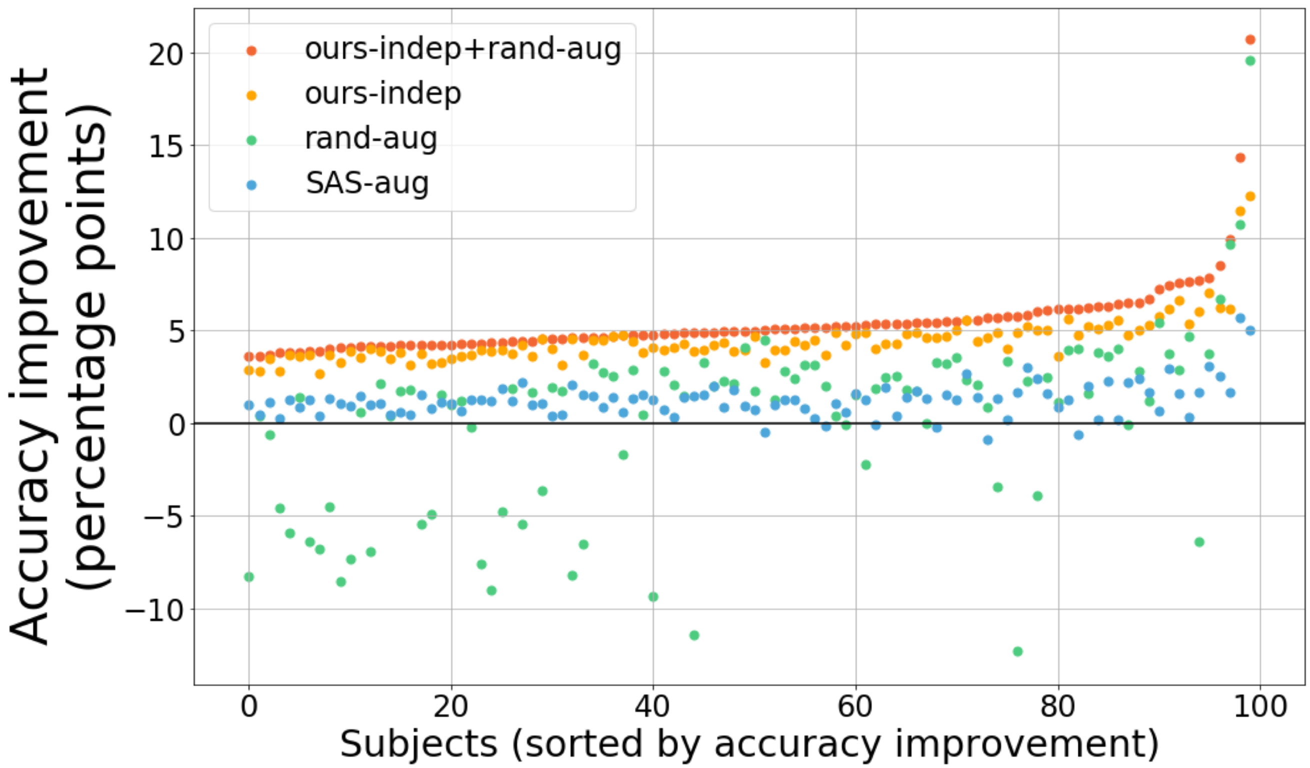
“Supervised” Network

- Train supervised segmentation on synthesized “realistic” data accuracy increase $76\ (\pm 14) \rightarrow 81.5\ (\pm\ 12)$



CVPR2019

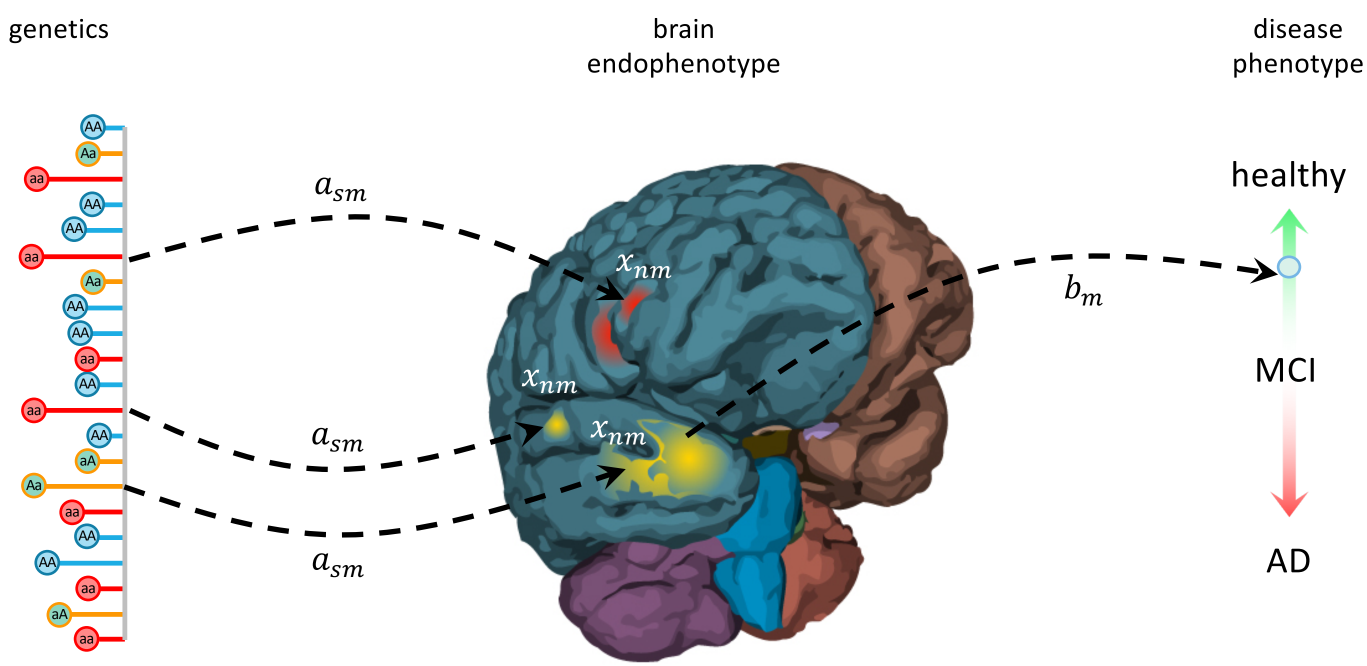
Results

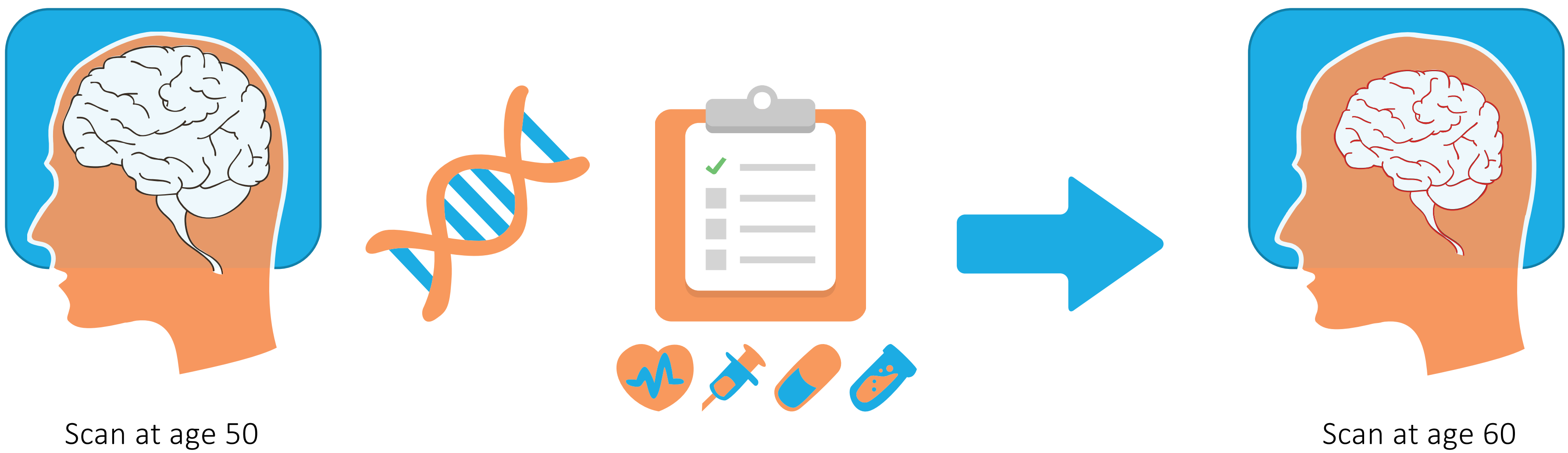


Zhao et al, in submission

Outline

- Overview of Medical Imaging
  - Utility and properties
- Example: Segmentation
  - *Classical* and deep learning approaches
- Example: Registration (alignment):
  - Optimization and learning approaches
- **Example: Imaging Genetics**
- Takeaways

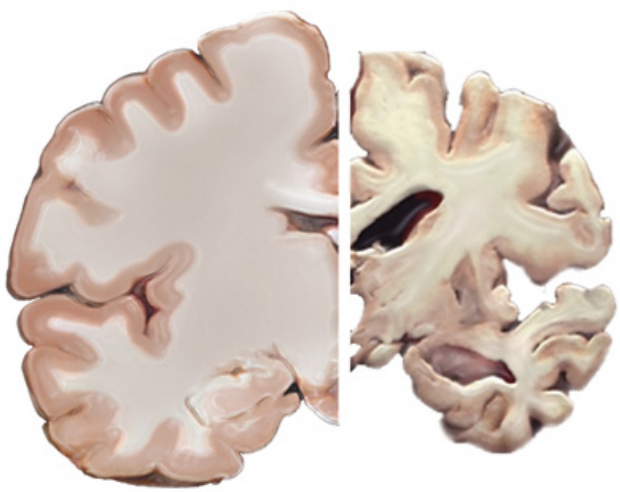

genetics
brain endophenotype
disease phenotype
a_{sm}
a_{sm}
a_{sm}
x_{nm}
x_{nm}
x_{nm}
b_m
healthy
MCI
AD
AA
Aa
aa
aA

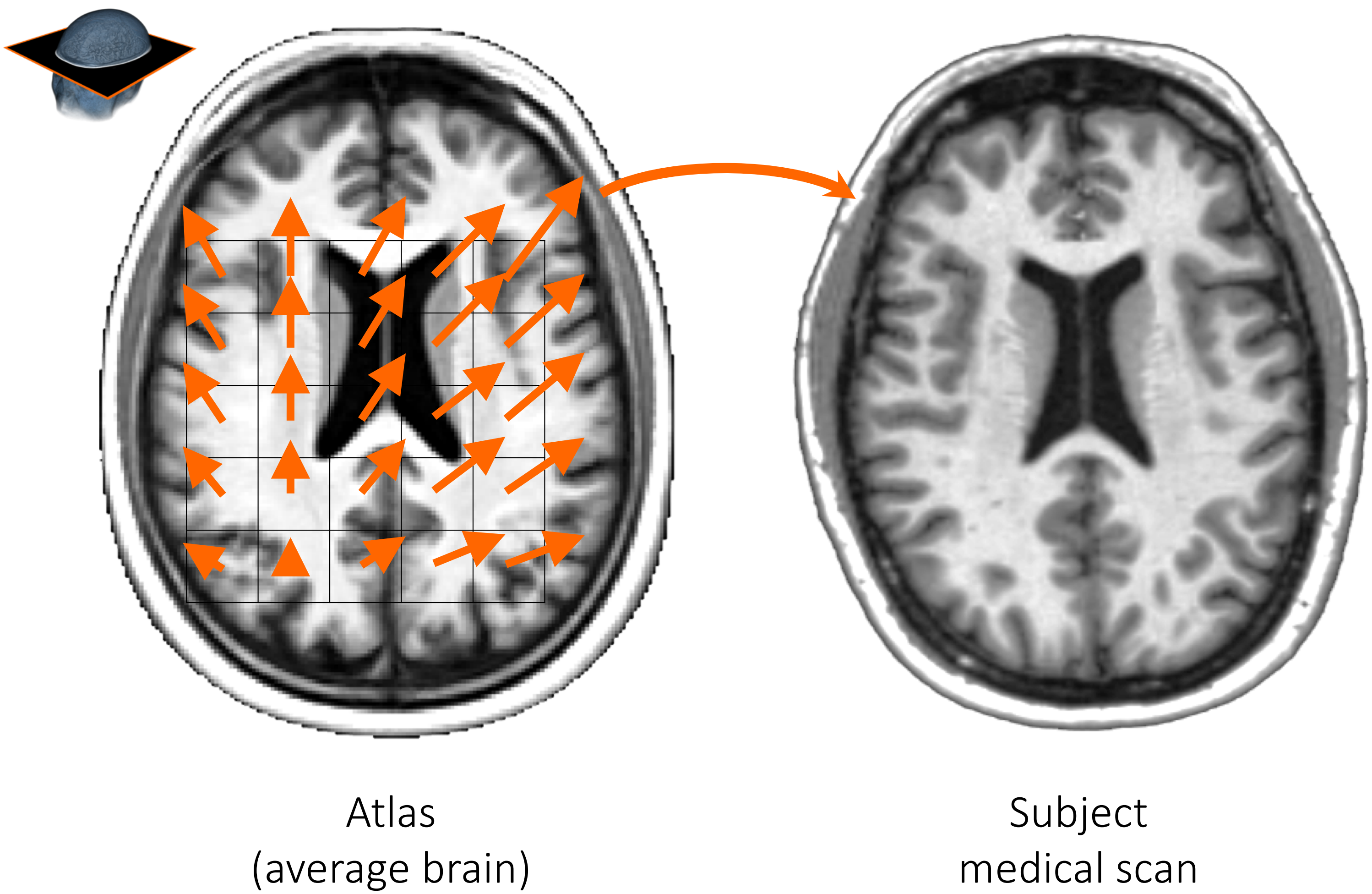

Scan at age 50
Scan at age 60

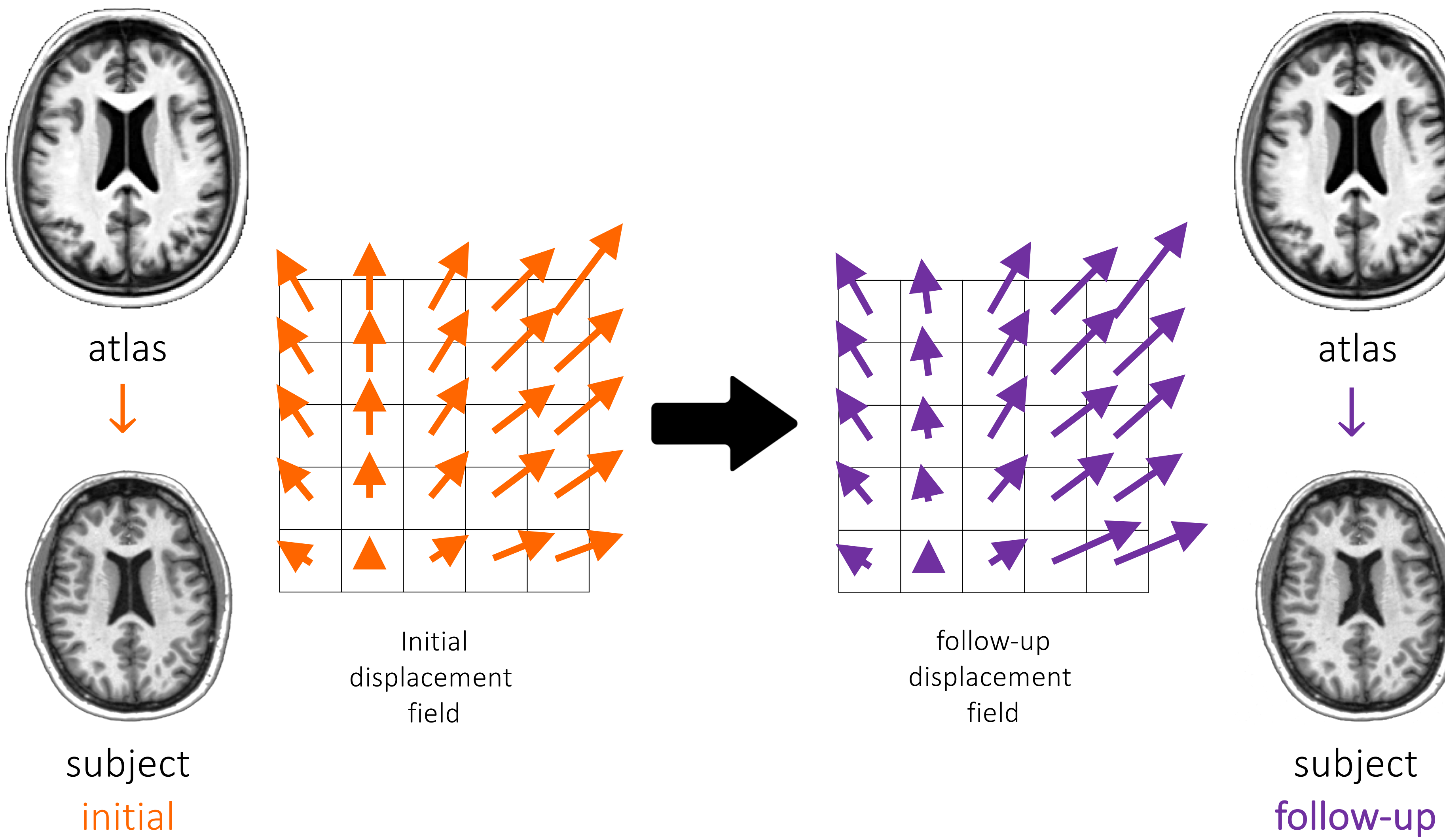
Motivation

Healthy Brain

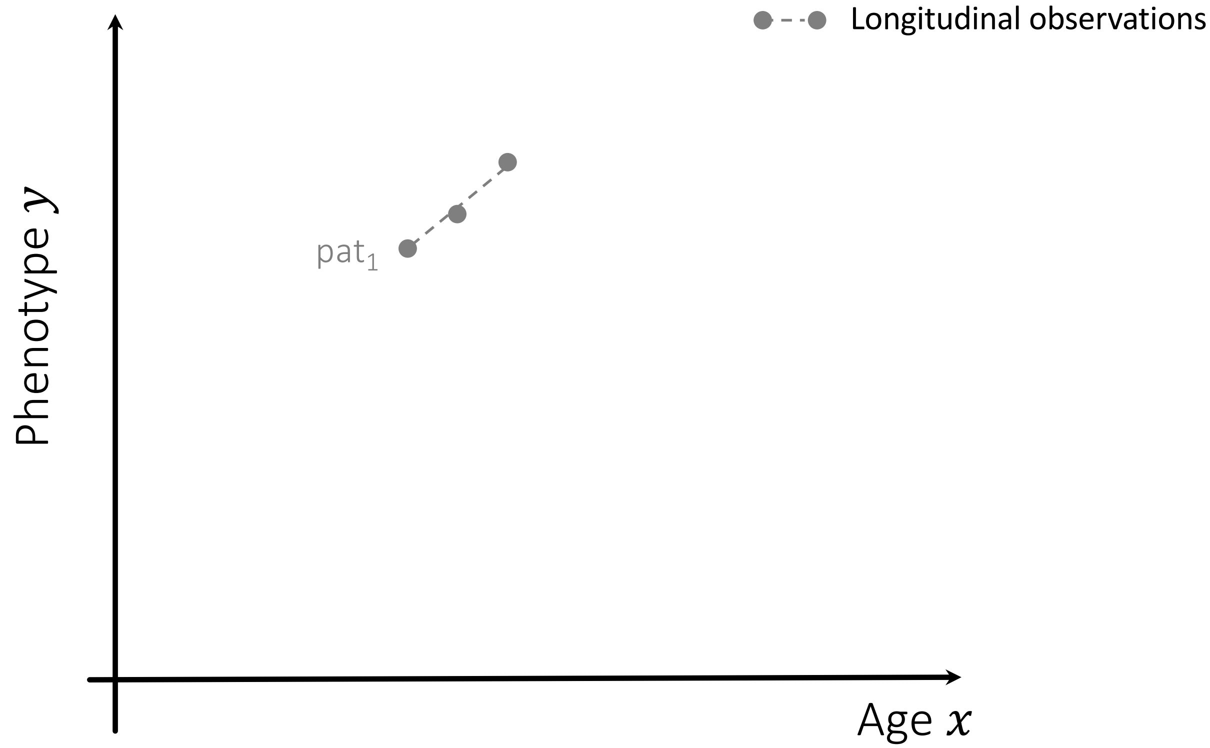
Severe Alzheimer's

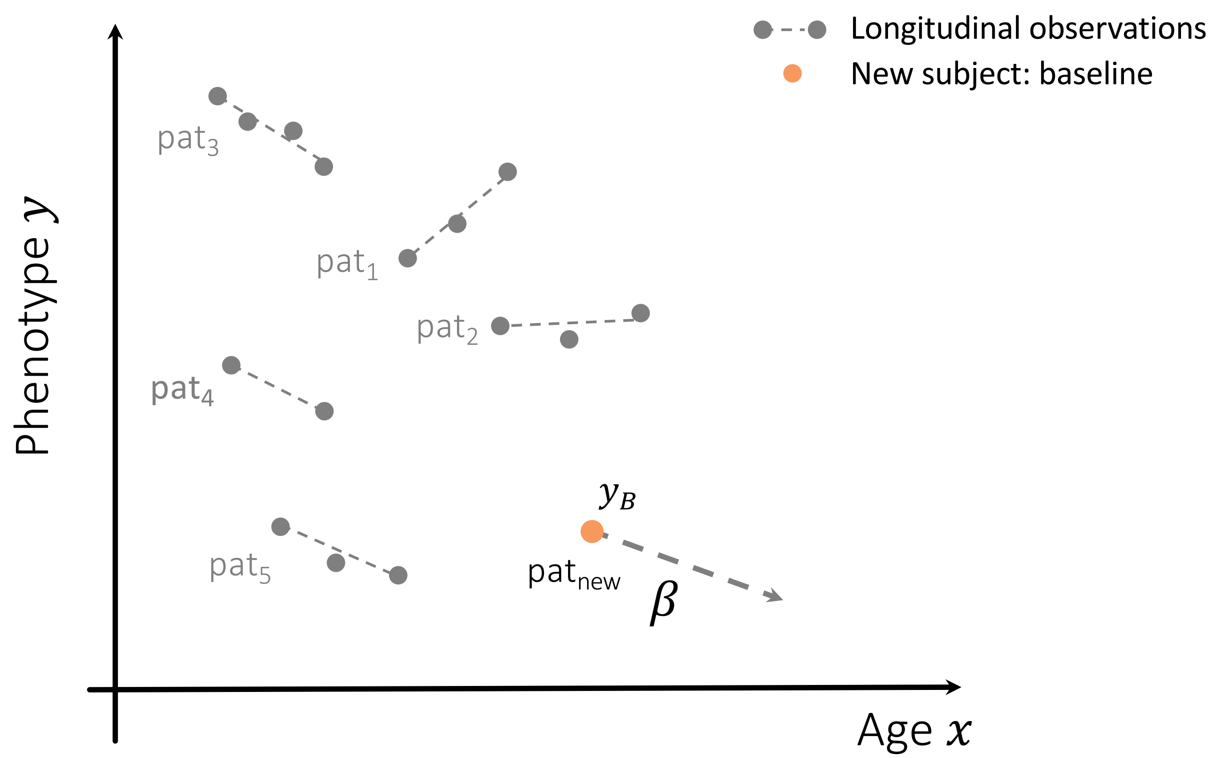



Atlas
(average brain)
Subject
medical scan

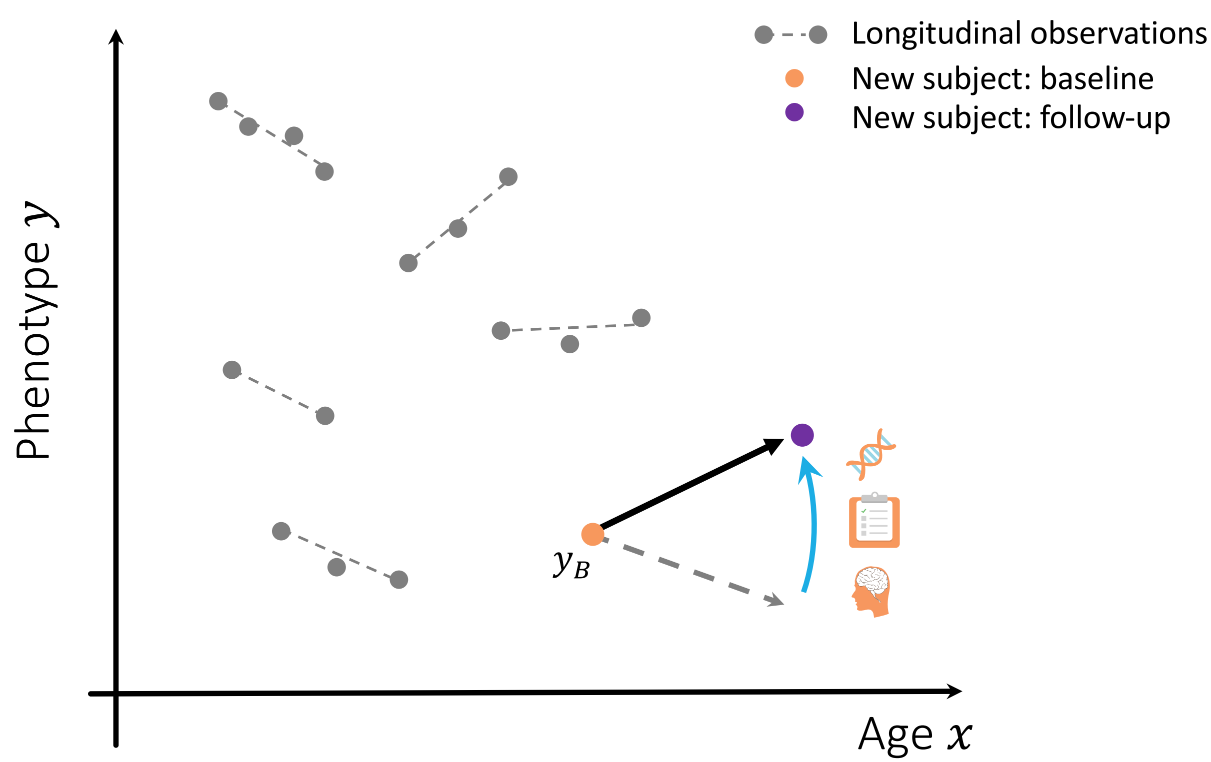

atlas
subject
initial
Initial
displacement
field
follow-up
displacement
field
atlas
subject
follow-up

Scalar Phenotype Change


Longitudinal observations
Phenotype y
pat$_1$
Age x



$$y_t = y_B + \Delta x_t \beta$$



$$y_t = y_{\mathrm{B}} + \Delta x\,(\beta + h_G + h_C + h_y) + \epsilon$$

Model

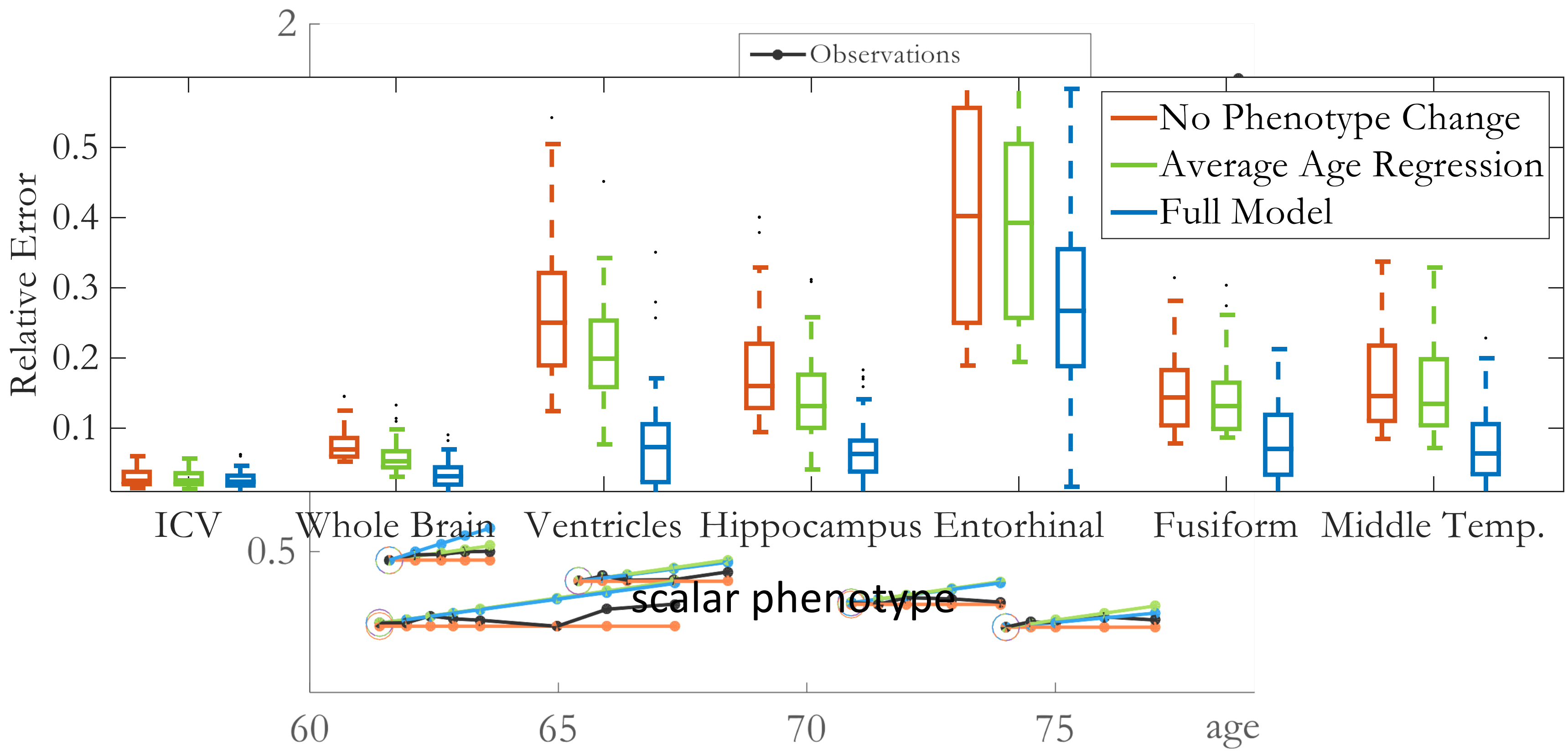
$$\boxed{y_t = y_{\mathrm{B}} + \Delta x\,(\beta + h_G + h_C + h_y) + \epsilon}$$

$$h_X(z_i) \sim \sum_j \overbrace{\alpha_j}^{\text{subject } j \text{ param.}} \underbrace{K_X(z_i, z_j)}_{\text{phenotype } z \text{ similarity}}$$

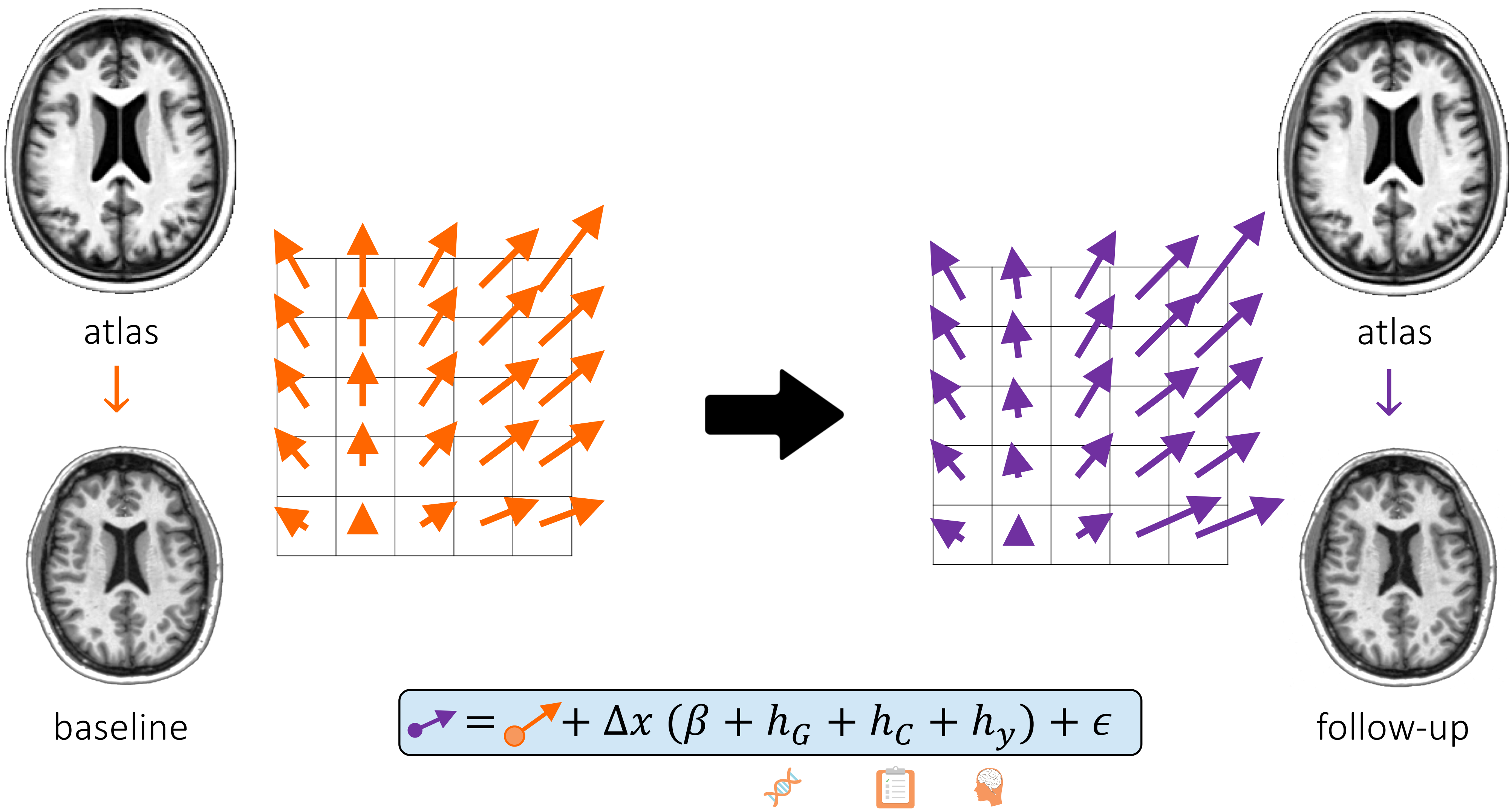
Genetics Kernel

$\boxed{\text{Main parameters: } \alpha_j \text{ and } \beta}$

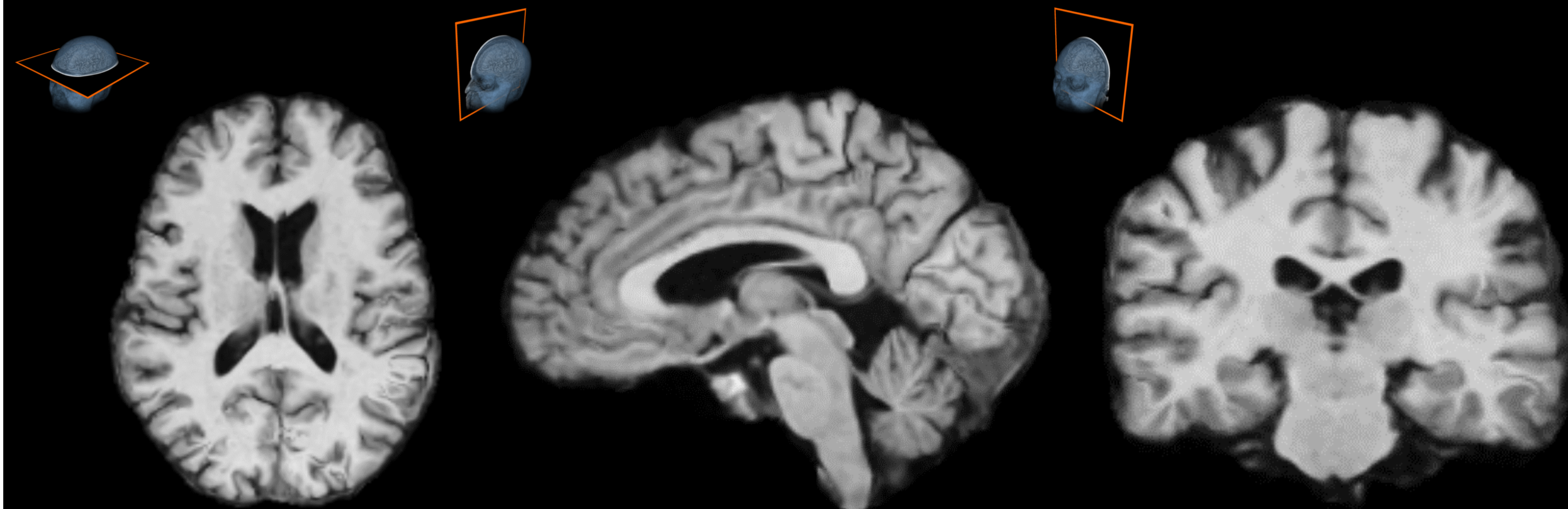
Scalar Phenotype Results

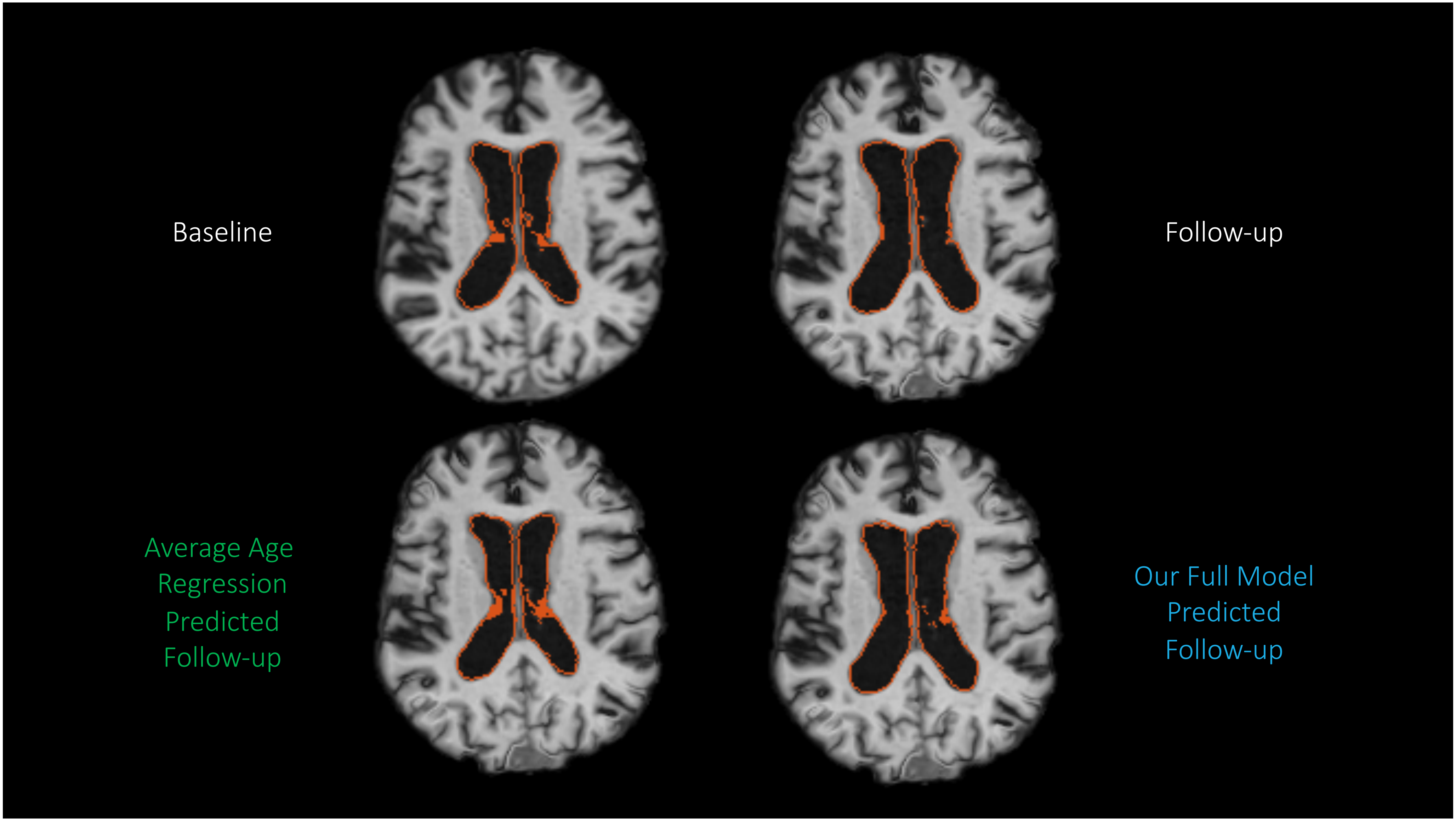


Predict Anatomical Scans

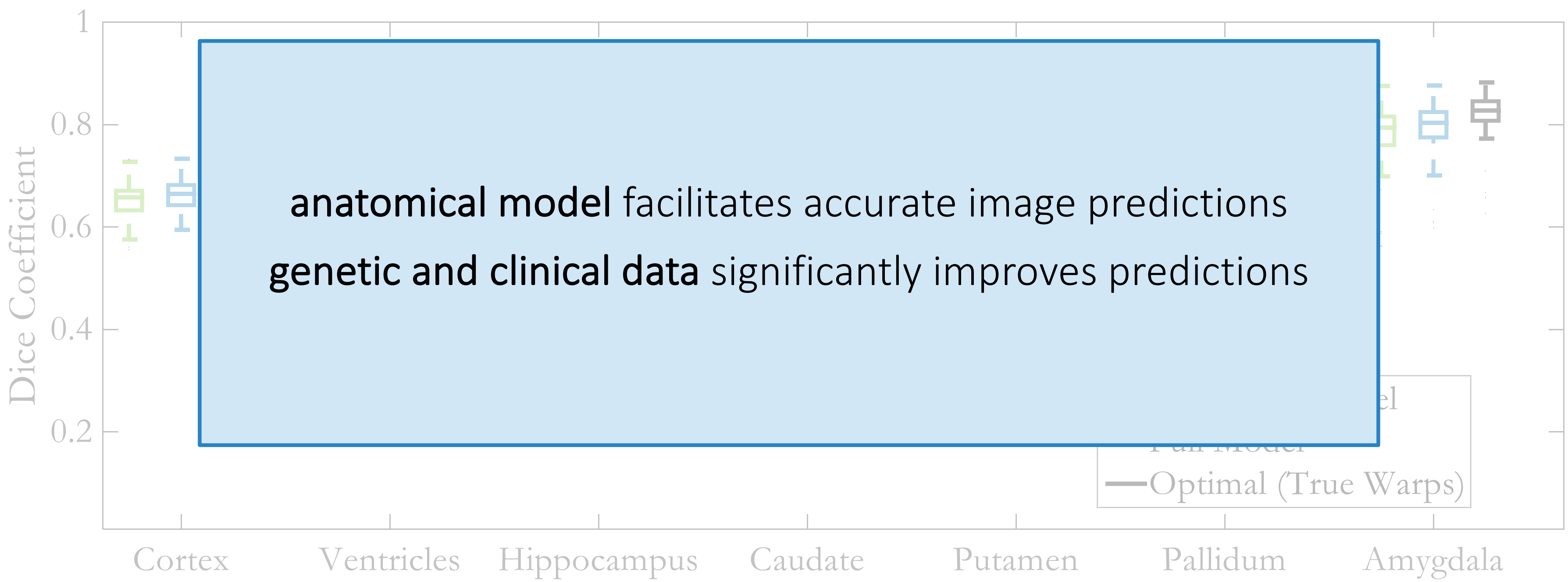

atlas
baseline
atlas
follow-up
$= \quad + \Delta x \, (\beta + h_G + h_C + h_y) + \epsilon$

Predicted follow-up scans




Baseline
Follow-up
Average Age
Regression
Predicted
Follow-up
Our Full Model
Predicted
Follow-up

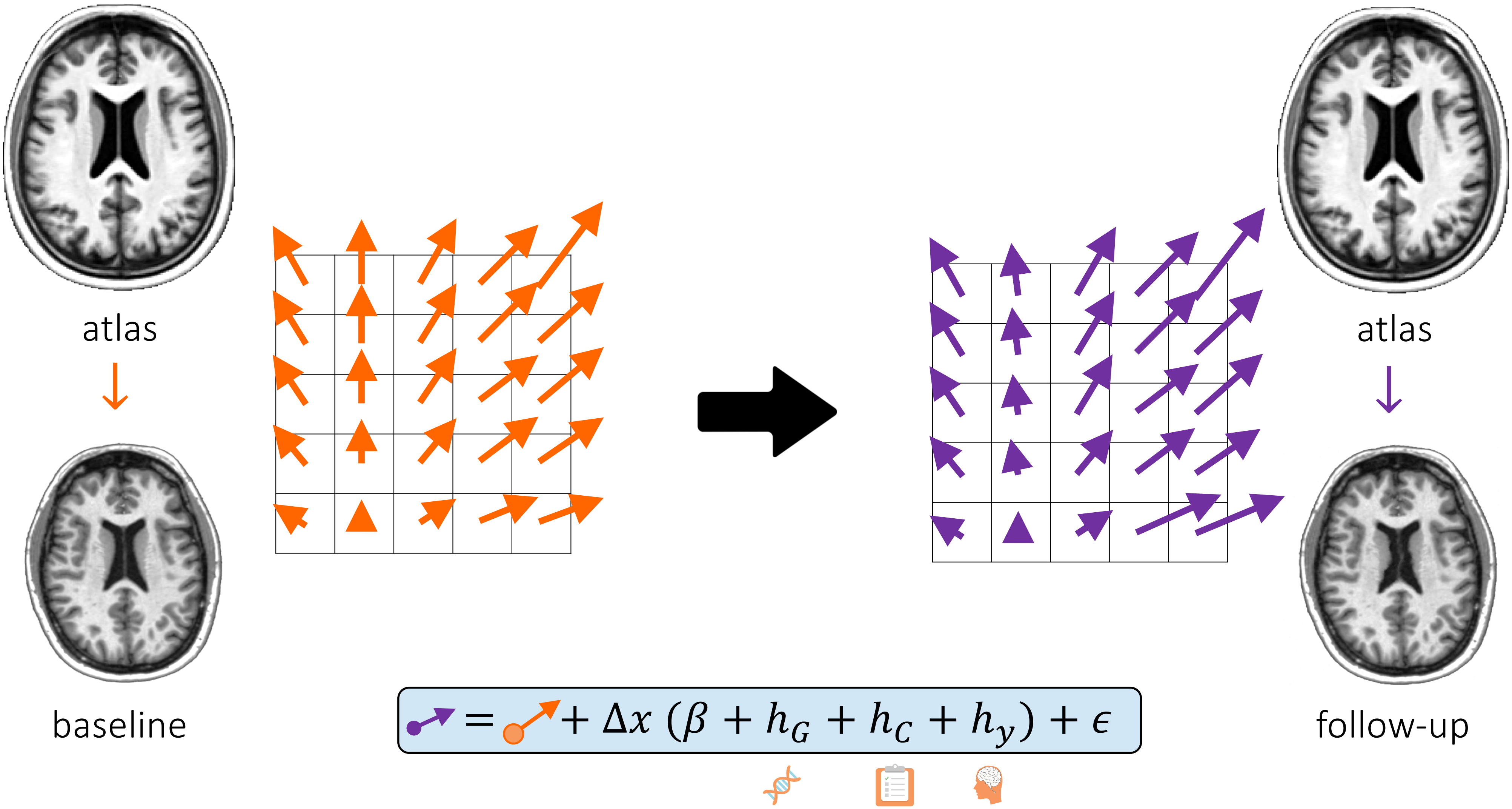
Volume overlap: predicted vs real structures



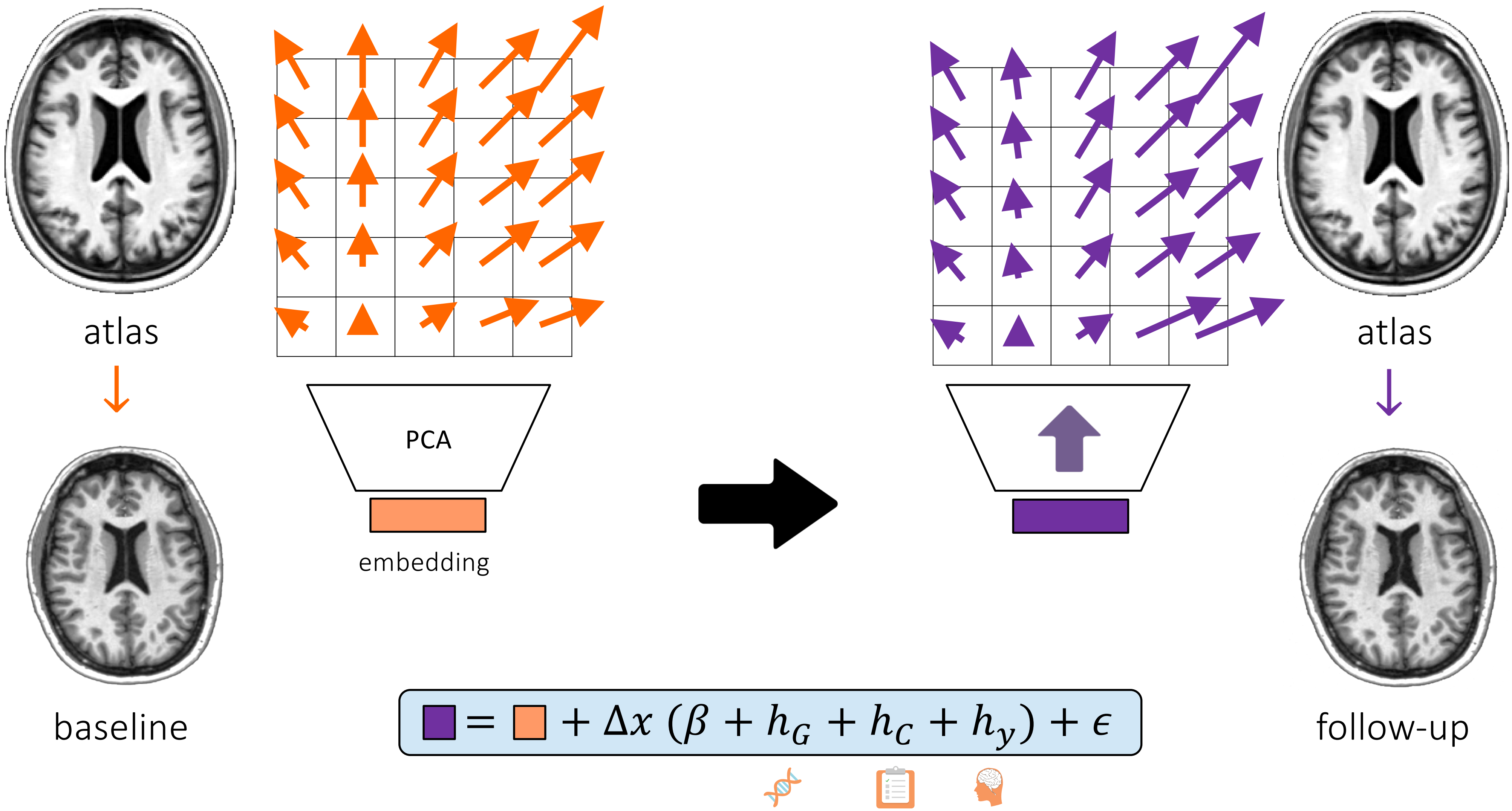
anatomical model facilitates accurate image predictions

genetic and clinical data significantly improves predictions

Questions?



atlas
baseline
atlas
follow-up
$= + \Delta x\,(\beta + h_G + h_C + h_y) + \epsilon$

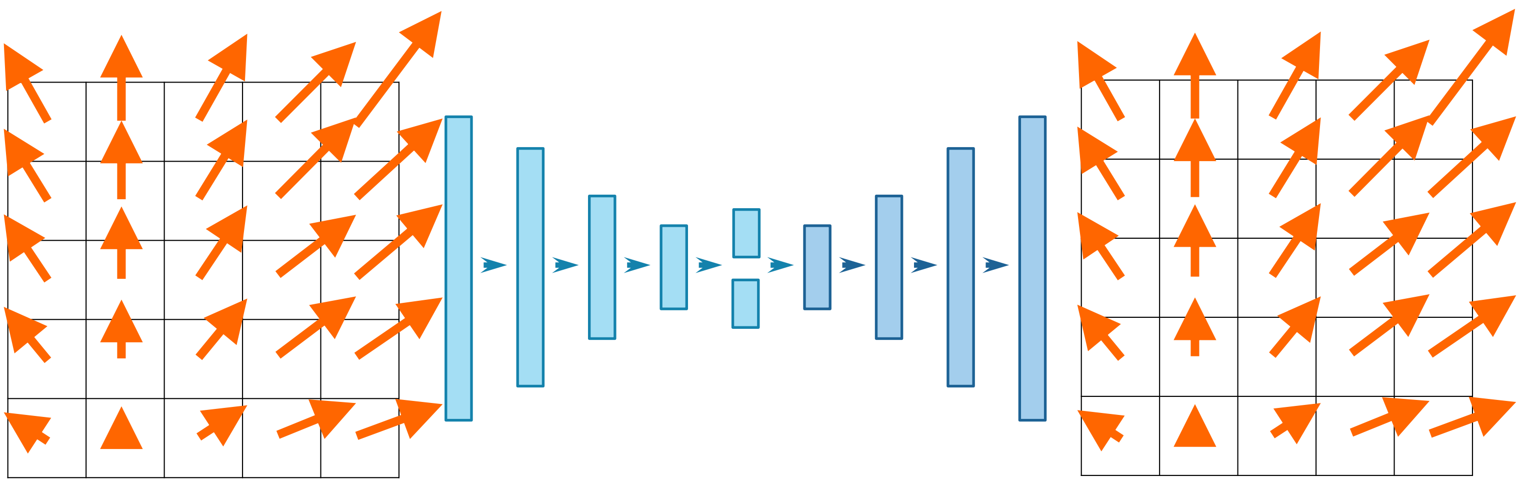

atlas
baseline
PCA
embedding
atlas
follow-up
$\blacksquare = \square + \Delta x\,(\beta + h_G + h_C + h_y) + \epsilon$

Linear models don’t capture imaging or deformations well

What to do?

Gaussian Process Prior VAE

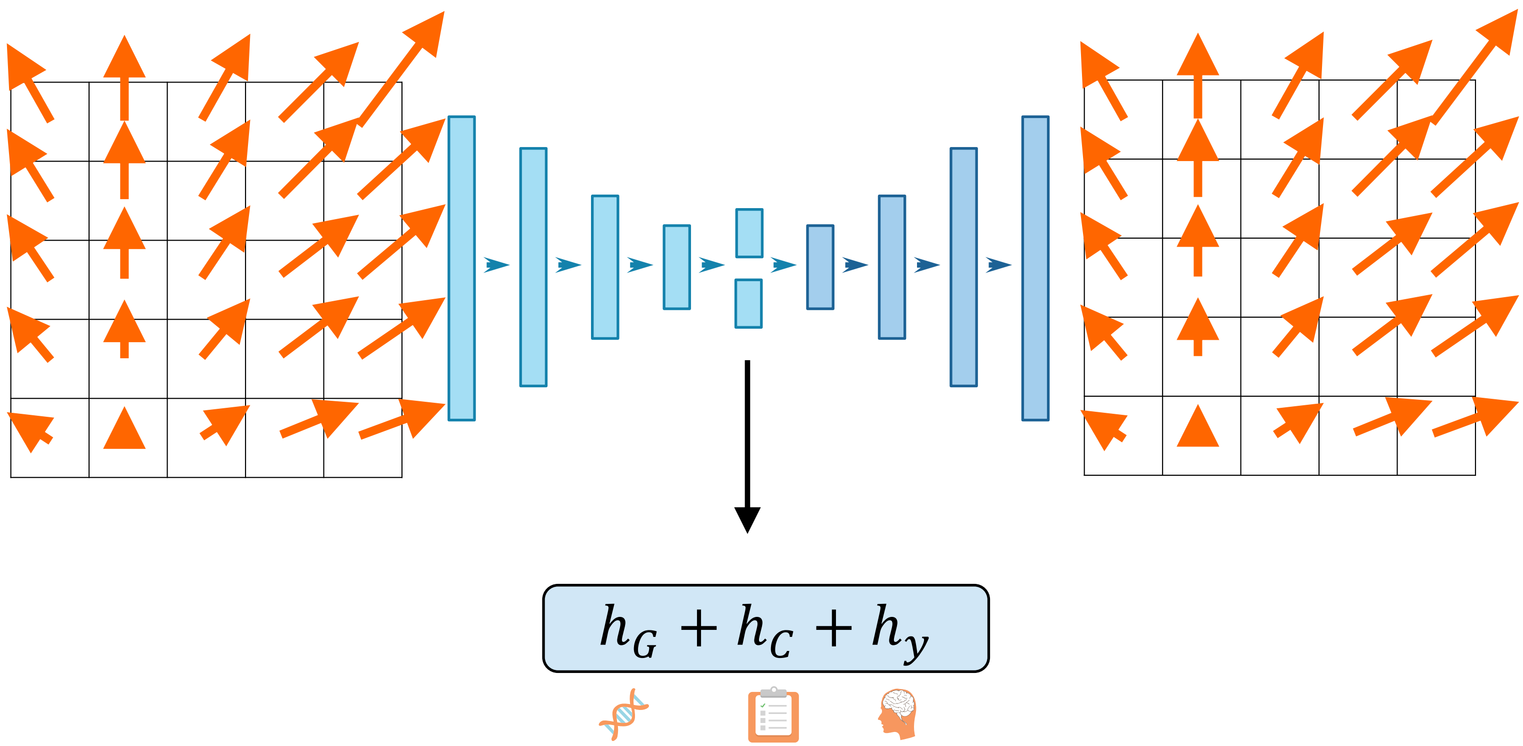
- Capture deformations with a VAE: $p(z)\, p(\phi|z)$



NeurIPS 2018

Gaussian Process Prior VAE

- Capture deformations with a VAE: $p(z)\, p(\phi|z)$
- Insert external data in the prior $p(z; G, C, y)$ using Gaussian process!



NeurIPS 2018

Conclusions

- A lot of medical imaging data
  - Machine (deep) learning enabling fast, successful methods
- In realistic scenarios, usually few **labelled** images
- Combine **learning** concepts and **clinical** knowledge
  - Limited supervised data: leverage unlabeled data
  - Large data: anatomically regularized deep networks
- Measure success if you impact **downstream clinical tasks**!

Outline

- Overview of Medical Imaging
  - Utility and properties
- Example: Segmentation
  - *Classical* and deep learning approaches
- Example: Registration (alignment):
  - Optimization and learning approaches
- Example: Imaging Genetics
- Takeaways

Takeaway Goals

- Problems
  - Help the clinicians or scientists (don’t replace them)
- Tools and approaches
  - Probabilities, convolutions, and anatomical models
  - Clinical interpretation
- Challenges
  - The systems don’t really work (yet)
- Opportunity
  - Impact healthcare (and research)!

Takeaway Goals

- Problems
  - Help the clinicians or scientists (don’t replace them)
- Tools and approaches
  - Probabilities, convolutions, and anatomical models
  - Clinical interpretation
- Challenges
  - The systems don’t really work (yet)
- Opportunity
  - Impact healthcare (and research)!