Understanding brain function with machine learning on large-scale data repositories

Bertrand Thirion, bertrand.thirion@inria.fr
Cognitive neuroscience

How are cognitive activities affected or controlled by neural circuits in the brain?
Cognitive neuroscience: From cognitive questions to data

Cognitive theories
- Sensory input
- Perceptual processing
- Working memory
- Consciousness
  - Cheater detection
  - Narrative mode of construal
  - Resource monitoring

Experimental paradigm

Scanner
- Brain mapping
- FMRI data

Stimuli
Cognitive neuroscience: From cognitive questions to data

Cognitive theories:
- Sensory input
- Perceptual processing
- Working memory
- Consciousness
  - Cheater detection
  - Narrative mode of construal
  - Resource monitoring

Experimental paradigm:
- Scanner
- Brain mapping
- FMRI data
- Experimental paradigm

Stimuli:
- Words
- Numbers
- Emotions
- Sounds
Cognitive neuroscience: Brain activity decoding

Brain activity decoding involves a cycle of encoding and decoding, where experimental paradigms are used to map the brain and FMRI data. Cognitive theories contribute to understanding the stimuli and processing involved.

- **Scanner**
- **Brain mapping**
- **FMRI data**
- **Experimental paradigm**
- **Encoding**
- **Decoding**
- **Cognitive theories**
  - Sensory input
  - Perceptual processing
  - Working memory
  - Consciousness
    - Cheater detection
    - Narrative mode of construal
    - Resource monitoring
  - Stimuli
    - Words
    - Numbers (35, 19)
    - Images

Nov 2022

Bertrand Thirion
The big data revolution is ongoing – in neuroimaging also!

The big data revolution is ongoing – in neuroimaging also!

Power failure: why small sample size undermines the reliability of neuroscience

Katherine S. Button¹,², John P. A. Ioannidis⁵, Claire Mokrysz¹, Brian A. Nosek⁴, Jonathan Flint⁵, Emma S. J. Robinson⁶ and Marcus R. Munafò¹

Scanning the horizon: towards transparent and reproducible neuroimaging research

Russell A. Poldrack, Chris I. Baker, Joke Durone, Krzysztof J. Gorgolewski, Paul M. Matthews, Marcus R. Munafò, Thomas E. Nichols, Jean-Baptiste Poline, Edward Vul & Tal Yarkoni
Problem: generalization across studies

“You cannot play 20 questions with nature and win”

[Newell A. Visual information processing; 1973]
[Poldrack & Yarkoni, Annu Rev Psycho 2016]

• Joint analysis: Use large studies to inform small studies (transfer learning)
  – Principle: leverage joint representations across datasets

• Mega-analysis: find semantic commonalities across studies
  – Difficulty: what common vocabulary across studies?
Outline

- Learning good representations for brain images: unsupervised approach
- Learning good representations for brain images: supervised approach
- *In the wild* brain activity decoding
- Dealing with semantics and data labelling issues
Learning good representations for brain images: unsupervised approach
Discovering structure in fMRI data
Discovering structure in fMRI

Can be captured by dictionary learning / sparse coding

[Olshausen Nature 1996]

→ Use of sparse PCA
High-dimensional fMRI

- $n =$ number of samples, $10^2$ to $10^6$
- $p =$ number of voxels, $10^5$-$10^6$

“Has anyone on the list run group-wise analysis on the HCP resting state data, and if so what tools did you use?

I am having memory issues when running more than 10 subjects and I was wondering if anyone has a way of getting around the large memory requirements when concatenating in time.”
Factorizing high-dimensional data

• Human Connectome project $n=4.10^6$, $p=2.10^5$, 4TB of data

• Online dictionary learning [Mairal et al. ICML 2009]

• How to go faster?
  - Work on batches of images and voxels
    • Online method in both samples and feature dimensions

[Mensch et al. ICML 2016, IEEE TSP 2018]
Stochastic gradient approaches

\[ \alpha_t(D) = \text{argmin}_{A \in \mathbb{R}^{k \times n}} \| x_t - D_{t-1} \alpha_t \|^2_F + \lambda \Omega(\alpha_t) \]

\[ D_t = \text{argmin}_{D \in C} \sum_{i=1}^{t} \| x_i - D \alpha_i \|^2_F \]
Stochastic gradient approaches

http://amensch.fr/research/2016/06/10/modl.html

\[
\alpha_t(D) = \arg\min_{A \in \mathbb{R}^{k \times n}} \| x_t - D_{t-1} \alpha_t \|_F^2 + \lambda \Omega(\alpha_t)
\]

\[
D_t = \arg\min_{D \in C} \sum_{i=1}^{t} \| x_i - D \alpha_i \|_F^2
\]

\[
\alpha_t(D) = \arg\min_{A \in \mathbb{R}^{k \times n}} \| M_t(x_t - D_{t-1} \alpha_t) \|_F^2
\]

+ \frac{\lambda s}{p} \Omega(\alpha_t)
Stochastic gradient approaches

10-fold gain in CPU time without loss in accuracy

[Menisch et al. ICML 2016, IEEE TSP 2018]

Can be used for recommender systems
Brain atlases

Learning good representations for brain images: supervised approach
Predictive modeling across datasets

4TB resting-state data
HCP900

OpenfMRI
HCP
Camcan
Brainomics...

40000 task
fMRI contrast maps
into one model

(a) Aggregation from many fMRI studies

Predictive modeling across datasets


[DiFuMo Dadi et al. Nimg 2020]
Predictive modeling across datasets

Transfer learning

Decoding from voxels
Decoding from multi-study networks
Chance level

Task fMRI study
Decoding accuracy on test set
Multi-study acc. gain

Nov 2022
Bertrand Thirion
Transfer learning
Transfer learning

![Graph showing multi-study accuracy gain vs baseline balanced accuracy and number of train subjects.]

- Standard decoding from voxels
- Decoding from functional networks
- Decoding from multi-study task-optimized networks

Accuracy gain compared to baseline median

---

Nov 2022  Bertrand Thirion
Small studies benefit more than large studies

- Pinel et al. '07
- Papadopoulos-Orfanos '12
- CamCan AV (Shafto et al. '14)
- Cohen '09
Resulting atlas
Resulting atlas

Available at https://github.com/arthurmensch/cogspaces

[Mensch et al. PCB 2021]
In the wild brain activity decoding
Open-ended brain decoding

What is this brain doing?

Which regions are predictive of tasks containing a given term?

- **Multilabel** classification problem
  - more than one class may be associated with each sample
- Predict occurrence of frequent terms

**Data:** experimental condition images

**Target**

sentences  
calculation  
tone listening  
tone counting  
successful stop
Classification results

[Schwartz et al. NIPS 2013, Varoquaux et al. PCB 2018]
Discriminative patterns

[Schwartz et al. NIPS 2013, Varoquaux et al. PCB 2018.]
An image database

Task fMRI repository
[Gorgolewski et al. 2015]

Currently 48k independent usable fMRIs

[Poldrack 2011], knowledge-base

• concepts: cognitive activity/state (e.g. working memory)
• tasks: standard experiment to probe it (e.g. n-back task)
From multi-study to universal decoder
Results (naive approach)
Fixing labels

Problem:
synonyms, false negatives (missing annotations)

→ Simple rules to impute labels:
Results (2): Yes!
Open the box

Non-controversial case
Open the box

decoding > encoding
Open the box

[Menuet et al. Scientific Reports 2022]
Dealing with semantics and data labelling issues

... by mining the neuroscientific literature
Need curated annotations

- Current ontology incomplete
- Bigger limitation = lack of consistent vocabulary
  [Poldrack & Yarkoni, Annu Rev psycho 2016]
- How to get those?
Mining neuroimaging literature

• Neuroimaging observations often stored in text.
• e.g “[…] in the anterolateral temporal cortex, especially the temporal pole and inferior and middle temporal gyri”
• Objectives:
  - transform neuroimaging publications into brain maps
  - meta-analysis of text-only corpora
Neuroquery

Extract

Shape and texture provide cues to object identity, but when objects are explored using vision and somatosensory (tactile) inputs, shape information is encoded in the anterior part of the inferotemporal cortex (ITa) and, in the monkey, in areas of the inferior parietal lobule (IPL). Evidence indicates that the C4 is involved during both visual and tactile shape processing. Here we used functional magnetic resonance imaging (fMRI) to examine whether visual texture-selective areas are similarly activated when observed and palpated.

We observed significant haptic texture-selective fMRI responses in areas of the primary auditory cortex (A1) and areas of the posterior portion of the inferior parietal lobule (IPL) that are functional homologues of ITa. These areas are activated during visual texture discrimination, although areas of somatosensory homologues of A1 are not activated during visual and haptic texture perception, these areas appear to be spatially distinct and modality-specific.

Transform

Fit

"primary auditory cortex"
Empirical evaluation of representations

Learning statistical correspondences across the literature is more effective than relying on atlases!

[Dockès et al. MICCAI 2018]
Leveraging semantics for better encoding

Semantic structure → map concepts with few/no data
Neuroquery

credit: https://neuroquery.org

[Dockès et al. elife 2020]
Conclusion

- Large-p data bring challenges:
  - Computation cost
  - Difficulty of statistical inference
- Solutions: compression, subsampling, ensembling
- Finding commonalities across cognitive studies is hard
- Big data approach:
  - Extract weak signals from huge amounts of data
  - Common representation across datasets (bottleneck)

Image processing may not be the hard part!
From good ideas to good practices: software

- Machine learning in Python
- Machine learning for neuroimaging
  - Classification of (neuroimaging) data
  - Network analysis

http://nilearn.github.io

BSD, Python, OSS

MNE

MEG + EEG ANALYSIS & VISUALIZATION
Acknowledgements

Parietal/Mind
G. Varoquaux,
A. Gramfort,
P. Ciuciu,
D. Wassermann,
D. Engemann,
A. Manoel,
D. Chyzhyk
A.L. Grilo Pinho,
E. Dohmatob,
A. Mensch,
J.A. Chevalier,
A. Hoyos idrobo,
D. Bzdok,
J. Dockès,
K.Dadi,
P. Cerda,
C. Lazarus
D. La Rocca
G. Lemaitre
L. El Gueddari
O. Grisel
M. Massias
P. Ablin
H. Janati
J. Massich
C. Petitot
JJ Torres

Other collaborators
J. Salmon
S. Arlot
M. Lerasle
P. Neuvial
S. Dehaene
R. Poldrack
J. Haxby
C. F. Gorgolevski
T. Yarkoni
G. Gautier

( & thanks for the data):
S. Dehaene
R. Poldrack,
J. Haxby
C. F. Gorgolevski
T. Yarkoni
G. Gautier

Human Brain Project

université PAris-Saclay

ANR