Causal models of brain dynamics

Unsupervised learning of optogenetic experiments via deep generative models

Tyler Benster

Deisseroth and Druckmann Labs

Qualifying Exam
Neurosciences PhD Program
Stanford University

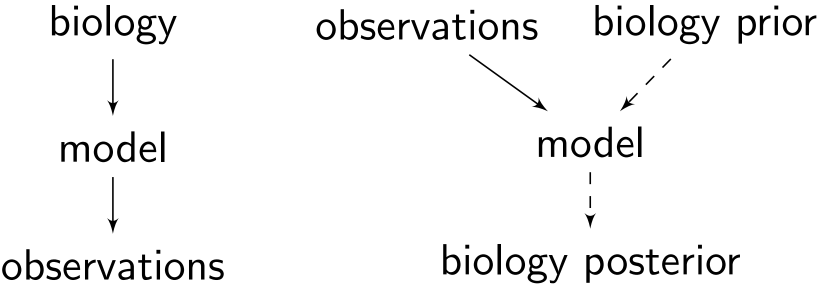
June 17, 2019

Roadmap

Whole-brain prediction facilitates construction of causal models

$$
\begin{tikzcd}
\text{biology} \arrow[d] \\
\text{model} \arrow[d] \\
\text{observations}
\end{tikzcd}
$$

Whole-brain prediction facilitates construction of causal models

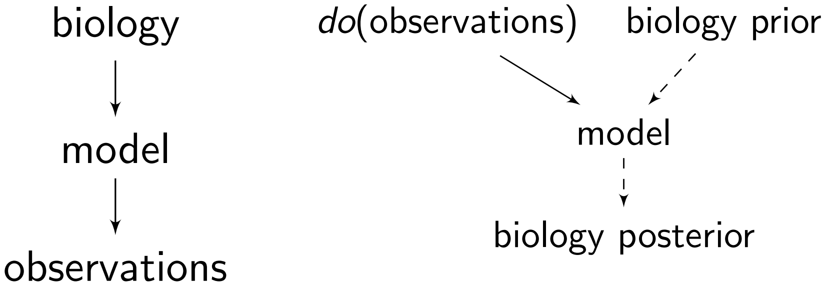


Whole-brain prediction facilitates construction of causal models

$$
\begin{tikzcd}
\text{biology} \arrow[d] \\
\text{model} \arrow[d] \\
\text{observations}
\end{tikzcd}
$$

$$
\begin{tikzcd}
\mathit{do}(\text{observations}) \arrow[dr] & & \text{biology prior} \arrow[dl, dashed] \\
& \text{model} \arrow[d, dashed] & \\
& \text{biology posterior} &
\end{tikzcd}
$$

Whole-brain prediction facilitates construction of causal models



- Observing whole-brain → fewer latent variables
- Interventions enable building causal model
- Causal models aid applications and interpretations

Aim 1: Spatial modeling with deep generative models

What is the most effective approach to predict whole-brain observations?

- State-of-art performance in video prediction is achieved by deep generative models
- Collected preliminary brain-wide calcium data during an optomotor behavior
- Initial modeling results suggest that spatial modeling out-performs traditional point process models of neurons

Aim 2: Optogenetic active learning

How do we resolve model underdetermination?

- For complex models fit to finite observations, multiple choices of parameters may perform equally well
- We can resolve this by testing if model substructures are causal with optogenetics
- First experiment failed due to poor optogenetic activation, but a new more-sensitive opsin will make the experiment easier

Aim 3: Contribution of cell-types and circuit motifs

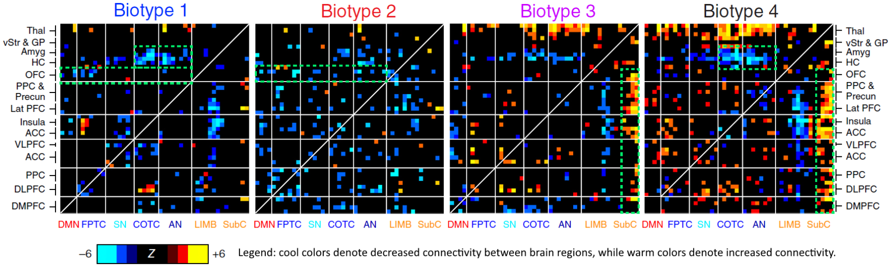
Do model substructures map to underlying biology?

- *in situ* hybridization (ISH) and connectome data contribute to understanding of colored graphs that underlie functional observations
- First attempt to add excitatory and inhibitory staining as an additional model input modestly hurt performance
- Training model to predict ISH data will force model to maintain representation of cell type
- Known circuit motifs can be used as a prior, or we can try to discover structure using structure active learning

Model-based summaries of health vs disease

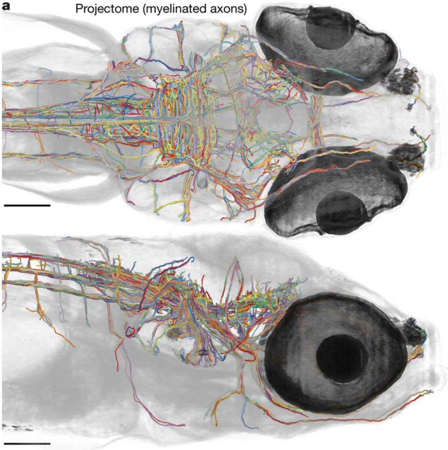
flickr:zigazou76 CC BY 2.0

Model-based summaries of health vs disease



Drysdale, Grosenick, et al. *Nature Medicine*. 2016

Challenge: high underlying complexity
as shown by partial zebrafish projectome

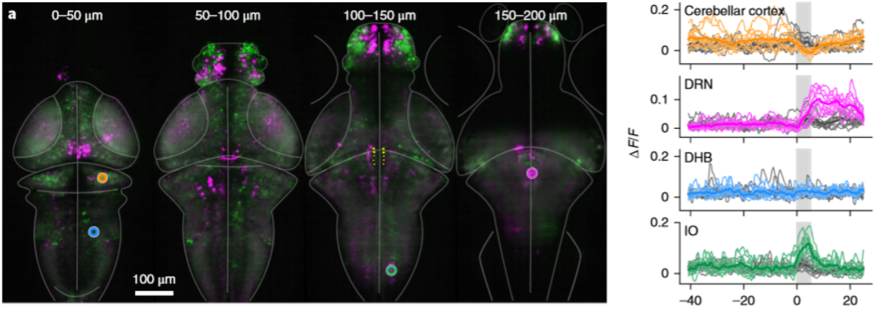


$\sim 80,000$ neurons at 7 days post fertilization (dpf)

Hildebrand, Cicconet, et al. *Nature* 2017
Hill, Howard, et al. 2003

Opportunity: massive datasets

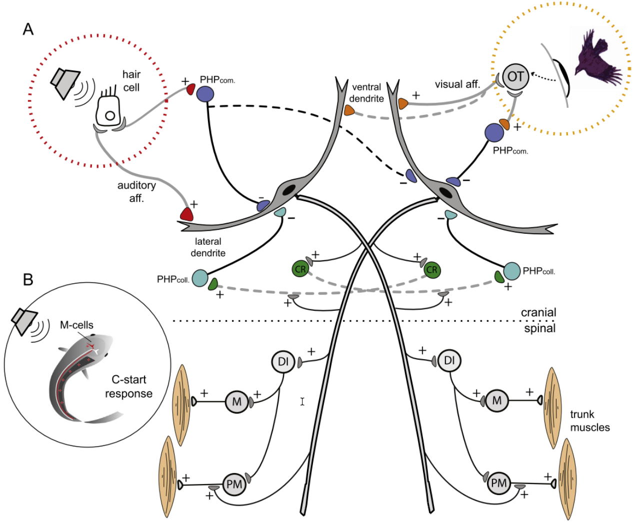
Whole-brain imaging & stimulation



left: increase in activity (green), decrease in activity (magenta), optogenetic stimulation sites (yellow) **right:** dorsal raphe nucleus (DRN), dorsal hindbrain (DHB), and inferior olive (IO)

Vladimirov, Wang, et al. 2018

Validate modeling results through comparison to well-characterized Mauthner circuit



Medan & Preuss, 2014

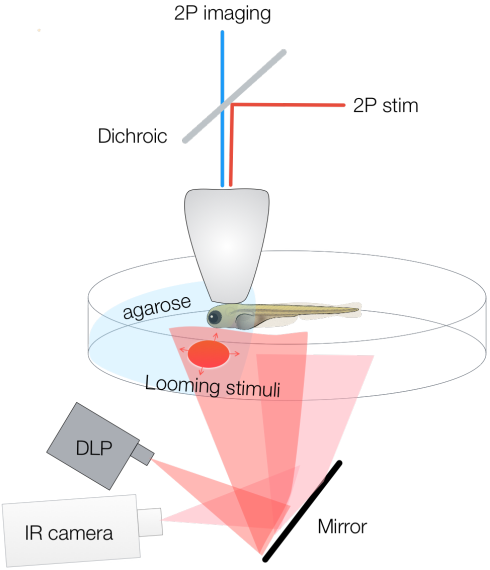
Aim 1: Spatial modeling with deep generative models

What is the most effective approach to predict whole-brain observations?

Hypothesis: deep learning spatial models will outperform point process models

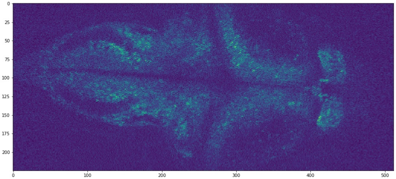
- Extract neuron fluorescent traces $\rightarrow$ build RNN
- Raw fluorescent observations $\rightarrow$ convolutional deep learning model
- Compare prediction performance on withheld test data

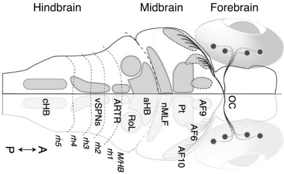
2P Experimental setup



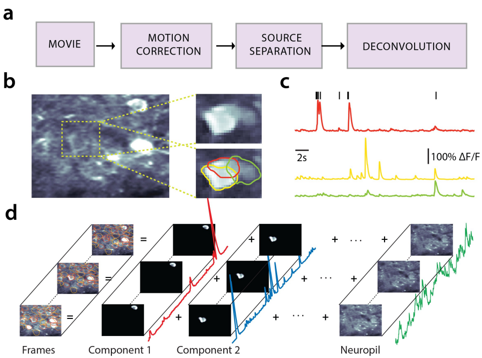
Whole-brain 2P calcium imaging

Z-projection of 19 planes, 4x real-time, 2Hz





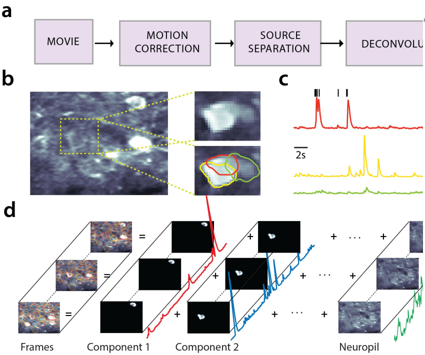
Current approaches to brain-wide modeling

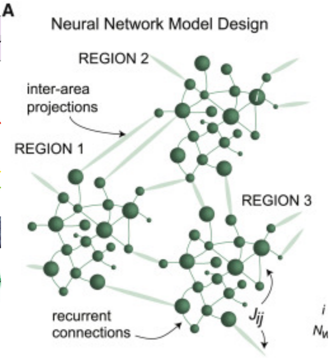


Extract neuron traces

Giovannucci et al. 2019.

Current approaches to brain-wide modeling



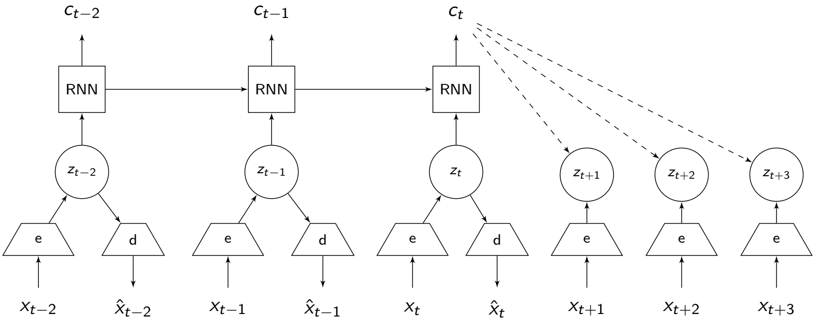


Extract neuron traces

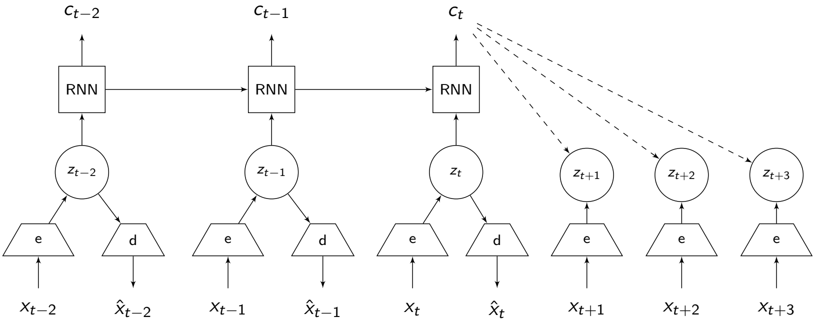
Add back spatial information

Giovannucci et al. 2019.
Andalman et al. 2019.

Latent-space volume prediction
Stochastic embedding



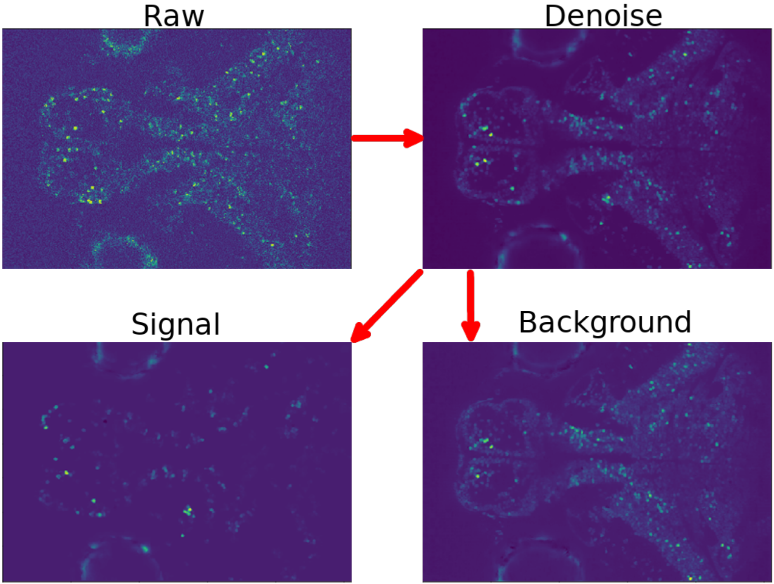
Latent-space volume prediction
Stochastic embedding



vs.

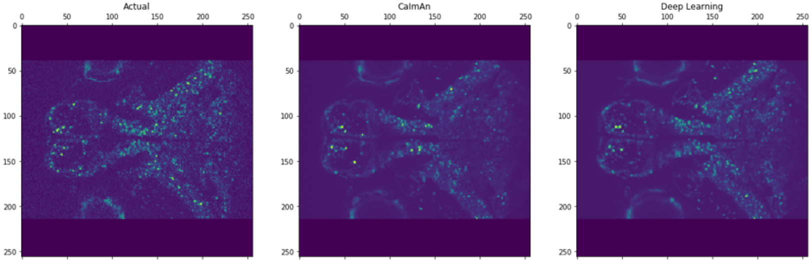
$$n_{t+5} = An_t$$

Mapping CNMF to space



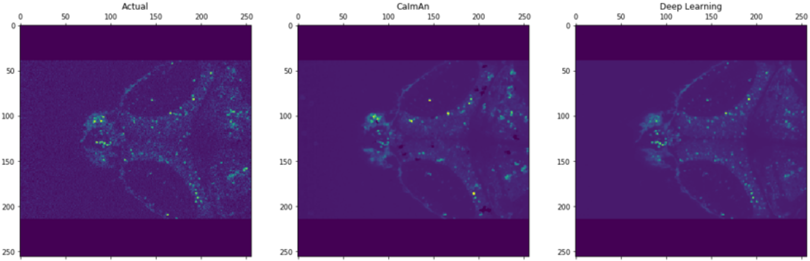
Train data: LS and VP perform equally well

$<$ 5% difference in MSE



Actual (left), least squares prediction (middle), volume prediction (right)

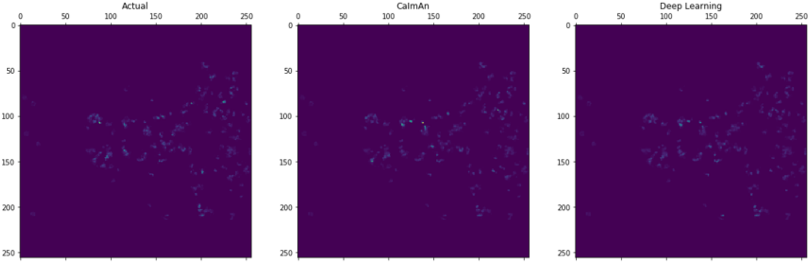
Test data: VP performs better
LS has 150% greater MSE than VP



Actual (left), least squares prediction (middle), volume prediction (right)

Masked test data: VP performs better

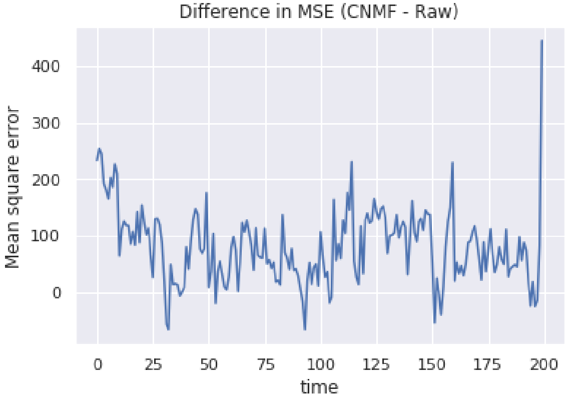
LS has 40% greater MSE than VP



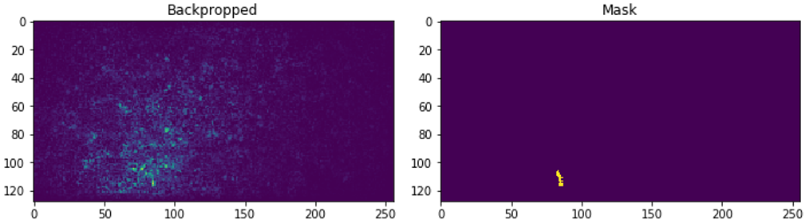
Actual (left), least squares prediction (middle), volume prediction (right)

CNMF preprocessing reduces VP performance

Evaluated loss on neuron mask



Extracting causal hypotheses
Voxels used for predicted locus coeruleus activation



Aim 2: Optogenetic active learning

How do we resolve model underdetermination?

Hypothesis: Model-based optimal experiment design will reduce underdetermination

- We can choose optimal optogenetic stimuli to maximally reduce model parameter uncertainty
- This can be potentially be done online
- We can evaluate how well we have done by attempting to track a brain trajectory

Active learning

Best to ask for category of which image?

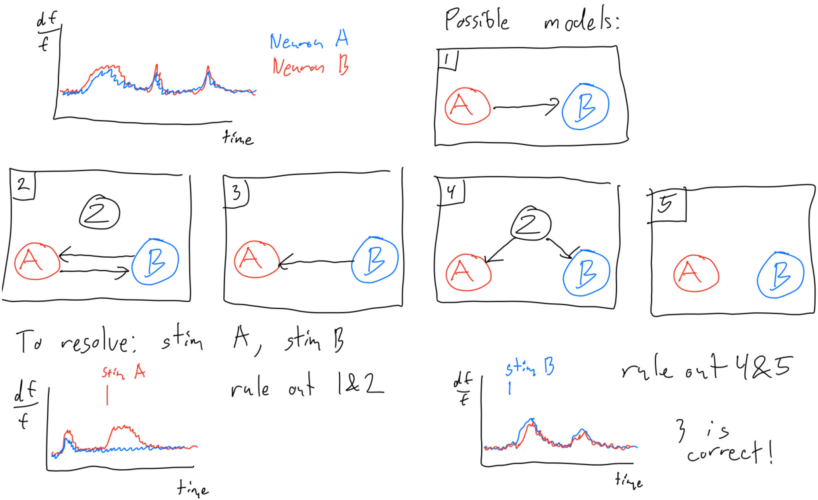
Best to ask for category of which image?



Cat

Mom & Dad. *personal correspondence*. 2016.
Instagram:atchoumthecat
Wikipedia CC BY-SA 3.0

Interventions resolve model underdetermination



Bayesian Active Learning by Disagreement

We want to maximize the decrease in expected posterior entropy of model parameters:

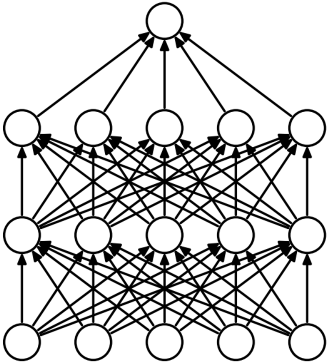
$$\underset{s_t}{\operatorname{argmax}}\, H[\theta|x_{1:t}] - \mathbf{E}_{x_{t+1}}[H[\theta|s_t, x_{t+1}, x_{1:t}]] \tag{1}$$

Entropy of model parameters θ is intractable. So we rearrange to:

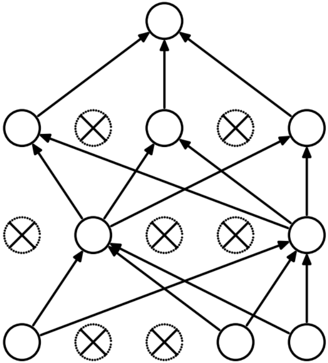
$$\underset{s_t}{\operatorname{argmax}}\, H[x_{t+1}|s_t, x_{1:t}] - \mathbf{E}_{\theta}[H[x_{t+1}|s_t, x_{1:t}, \theta]] \tag{2}$$

Houlsby, Huszár, et al. 2011

Bayesian deep learning via dropout



(a) Standard Neural Net



(b) After applying dropout.

Srivastava, Hinton, et al. 2014
Gal 2016

Collect data then iterate the approach offline

Data collection:

1. 180 trials of looming stimuli (1 hour)
2. 360 trials of random single-cell perturbation (1 hour)
3. 180 trials of looming stimuli (1 hour)

Data analysis:

- train on 80% of trials from [1 & 3]
- choose 60 trials from [2]
- Test on witheld trials from [1 & 3]

How much better can we do by choosing trials vs random trials in terms of test performance?

Online optimal experiment design

Data collection:

1. 225 trials of looming stimuli (1 hour 15 min)
2. 360 trials, model chooses each single-cell perturbation (1 hour)
3. 225 trials of looming stimuli (1 hour 15 min)

How well can we predict test data by training on 2 hours of looming stimuli vs 1 hour looming & 1 hour optogenetics?

Brain state replay

Data collection:

1. Acquire brain trajectory of interest
2. choose each stimulation pattern sequentially during resting state / experiment of interest
3. Stim brain to keep observations in line with [1]

How well can we track a previously observed trajectory?

Aim 3: Contribution of cell-types and circuit motifs

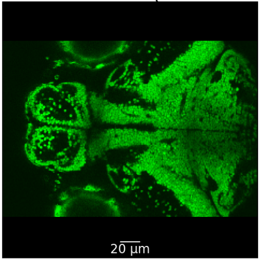
Do model substructures map to underlying biology?

Hypothesis: Enforcing causal biological constraints will improve model performance while aiding interpretatation

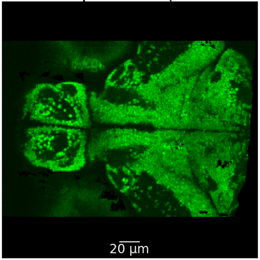
1. Requiring model to predict *in situ* hybridization allows for interpretation of cell-type contribution to dynamics
2. Template-based representations allow for unsupervised learning of purported circuit motifs

in situ cell type identification

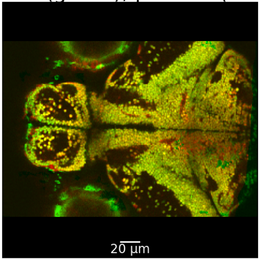
GCamP6s alive (isosbestic)



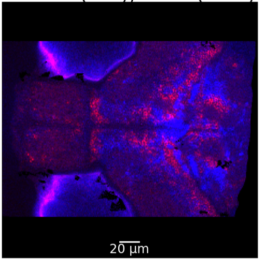
GCamP6s post-fix (isosbestic)

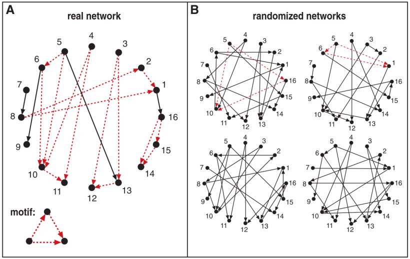


alive (green), post-fix (red)

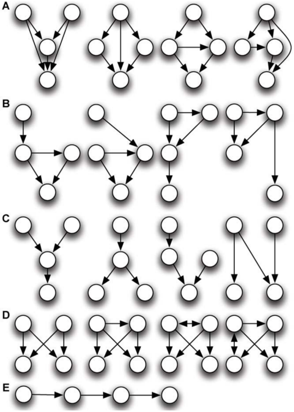


VGLUT (red), GAD (blue)





Schematic illustrating an over-represented motif



Over-represented motifs from *C. elegans* connectome

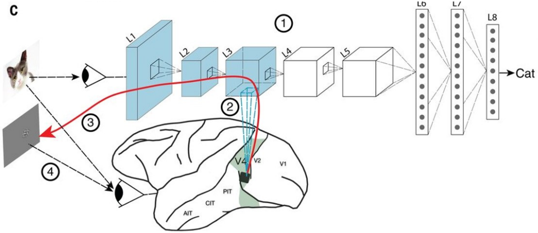
Milo et al 2002
Qian et al 2011

Conclusion

- Aim 1: Spatial modeling with deep generative models
  - Spatial modeling outperforms point process modeling in prediction accuracy
  - Next: repeat experiments and analyze tail movement prediction
- Aim 2: Optogenetic active learning
  - Established theoretical foundation for selecting optimal optogenetic stimuli to reduce model uncertainty
  - Next: create transgenic and validate single cell activation
- Aim 3: Contribution of cell-types and circuit motifs
  - Acquired preliminary dataset with co-registered functional imaging and excititory & inhibitory staining
  - Next: use model to predict cell type and evaluate influence of cell type on dynamics

Stimulate retina → control v4

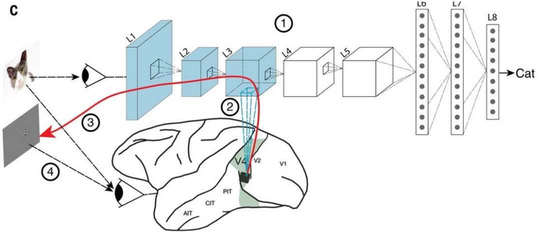
activate only one neuron of population with overlapping receptive fields

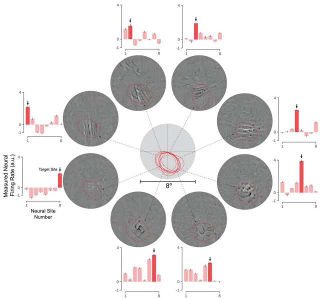


Bashivan, Kar & DiCarlo. *Science*. 2019

Stimulate retina → control v4

activate only one neuron of population with overlapping receptive fields

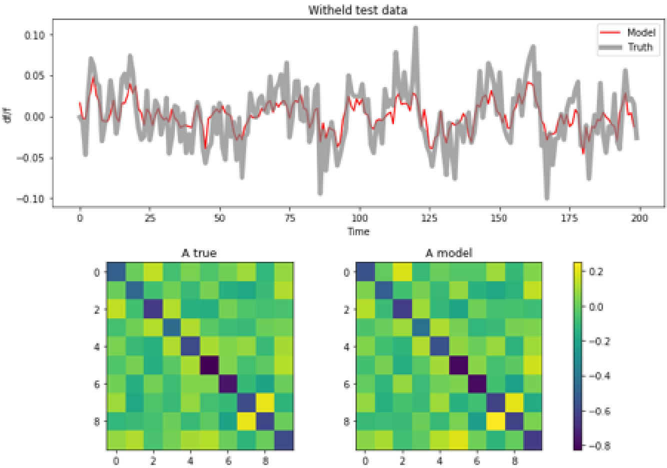




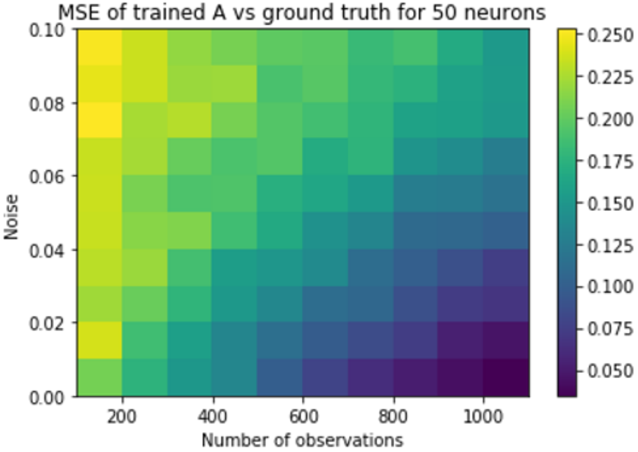
Bashivan, Kar & DiCarlo. *Science*. 2019

Linear Dynamical System with noise

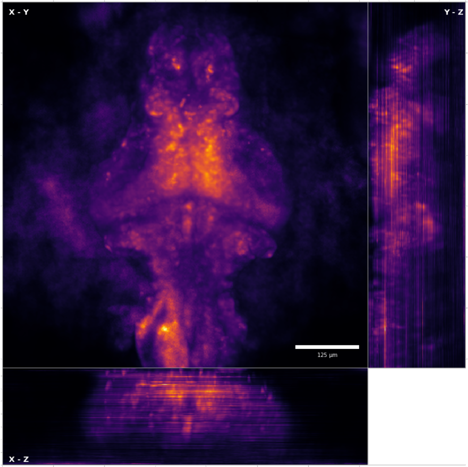
$$x_{t+1} = Ax_t + \epsilon$$



Reconstruction breaks down as noise increases



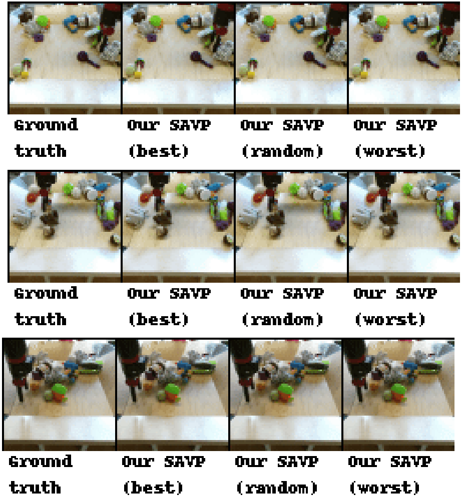
Whole-brain imaging at 200Hz
exemplified by preliminary Extended Light Field Microscopy reconstruction



1.4 TB/hr

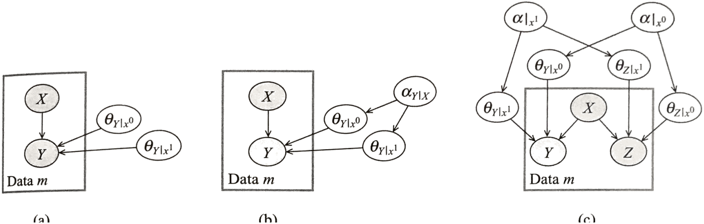
Noah Young. *unpublished*. 2019

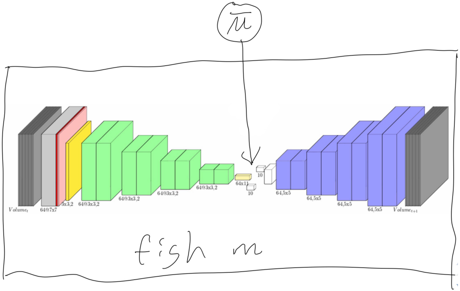
Learning physics from video prediction



Lee, Zhang, et al. 2018

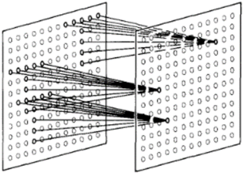
Extracting principles from multiple fish



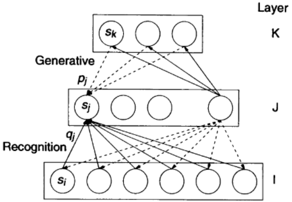


Koller & Friedman. *Probabilistic Graphical Models*. 2009.

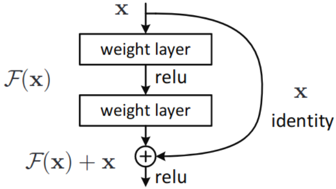
Modern deep learning toolkit



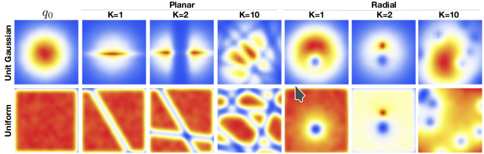
Convolutional neural networks[1]



Variational Bayesian learning[3]



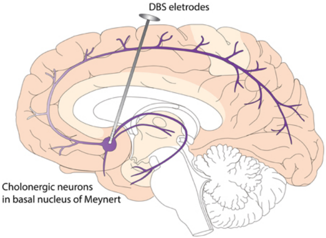
Deep residual structure [2]



Normalizing flows [4]

Fukushima 1980 [2]He et al 2015 [3]Hinton et al 1995 [4]Rezende&Shaker 2016

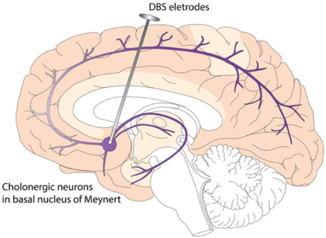
Brain-computer interfaces
"writing" into the brain



Zhang & Kim 2015

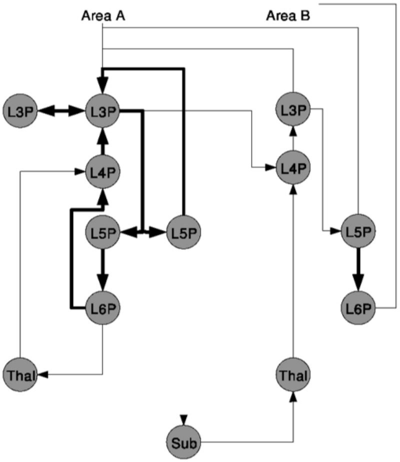
Brain-computer interfaces

"writing" into the brain



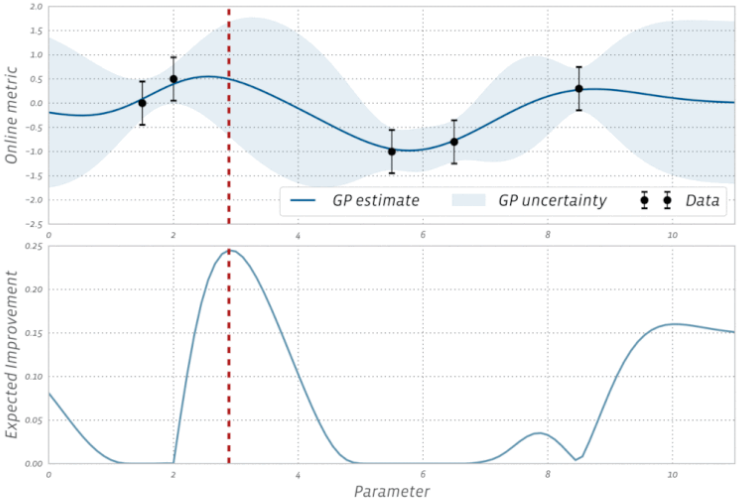
Zhang & Kim 2015 *The Matrix* 1999.

Canonical circuits



Bayesian optimization of latent space uncertainty

multi-neuron stim



github: Facebook/Ax