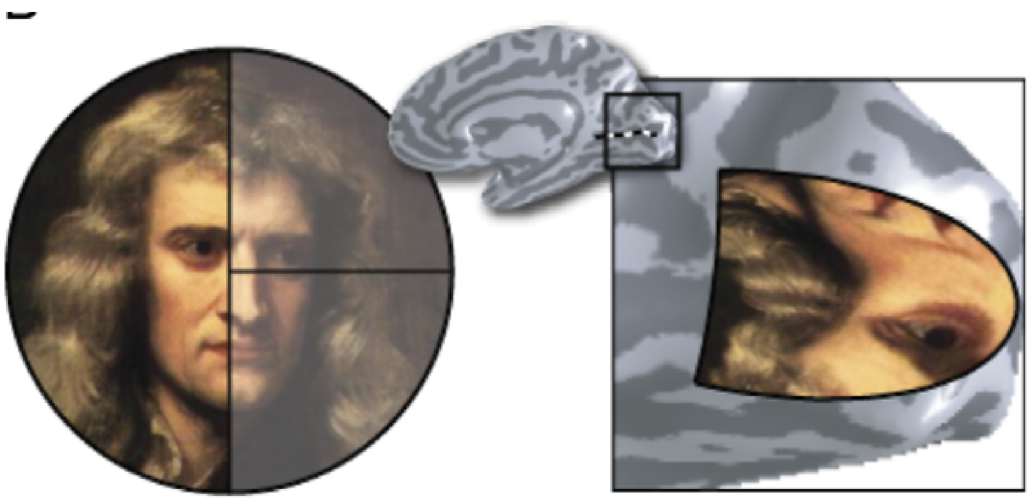
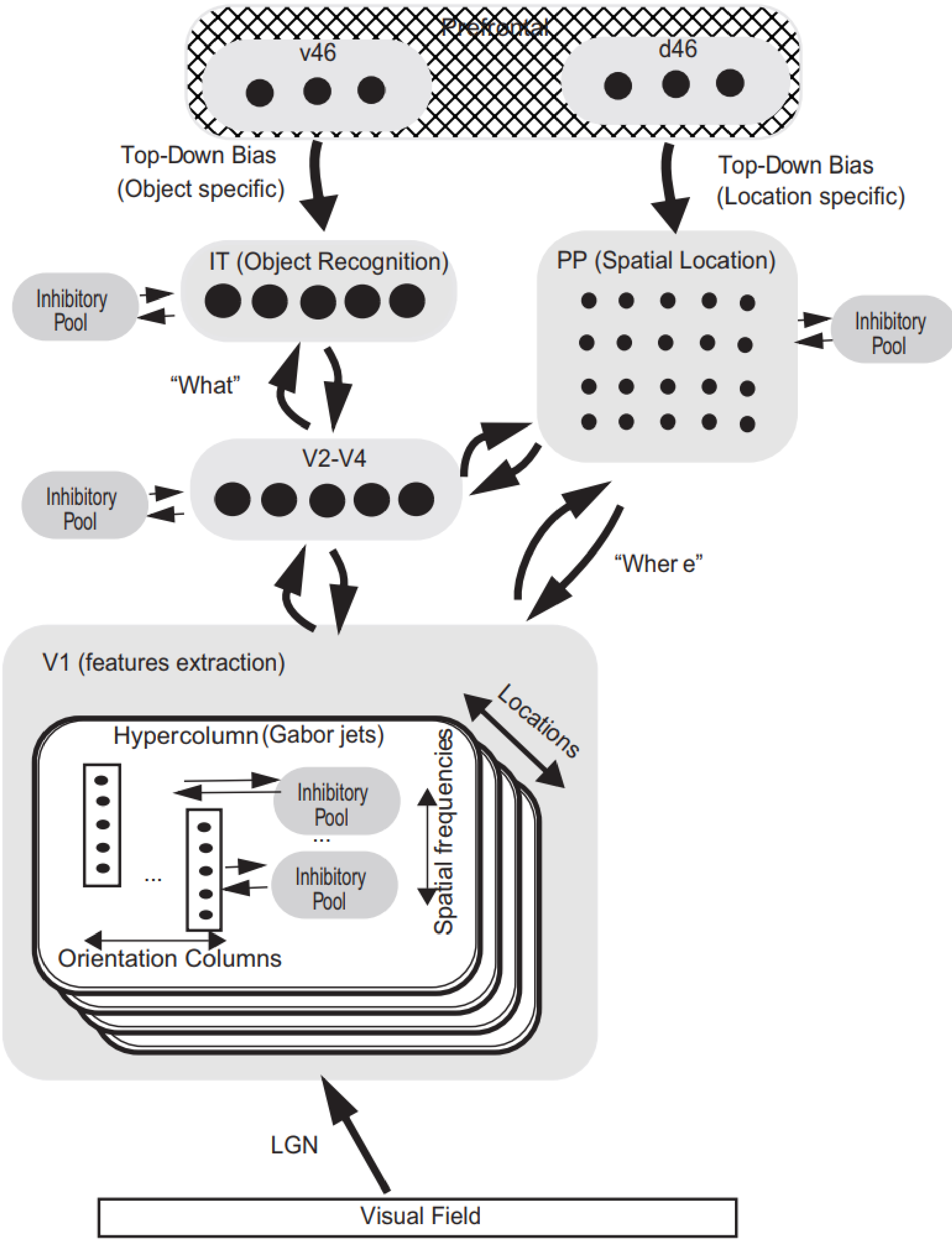
HCP Retinotopy Project Diary

GUIDES

- https://compneuro.neuromatch.io/projects/docs/project_templates.html#retinotopic-mapping-with-fmri
- https://compneuro.neuromatch.io/projects/docs/project_guidance.html



The systems-level architecture of a model of the cortical mechanisms of visual attention. The system is essentially composed of six modules structured such that they resemble the two known main visual pathways of the primate visual cortex. Information from the retino-geniculo-striate pathway enters the visual cortex through area V1 in the occipital lobe and proceeds into two processing streams. The occipital-temporal stream leads ventrally through V2–V4 and IT (inferotemporal), and is mainly concerned with object recognition. The occipito-parietal stream leads dorsally into PP (posterior parietal complex) and is responsible for maintaining a spatial map of an object's location.


Prefrontal
v46
d46
Top-Down Bias
(Object specific)
Top-Down Bias
(Location specific)
IT (Object Recognition)
PP (Spatial Location)
Inhibitory Pool
Inhibitory Pool
"What"
V2-V4
Inhibitory Pool
"Wher e"
V1 (features extraction)
Locations
Hypercolumn (Gabor jets)
Inhibitory Pool
...
Inhibitory Pool
...
Spatial frequencies
Orientation Columns
LGN
Visual Field

Project Resources:

Paper:

- Computational neuroimaging and population receptive fields [Fig. 1]: https://www.sciencedirect.com/science/article/pii/S1364661315000704
- Visual Field Maps in Human Cortex [check the “Traveling-Wave Method” section and “New Frontiers” section]:
- Population receptive field estimates in human visual cortex: https://www.sciencedirect.com/science/article/pii/S1053811907008269

GLM:

- https://compneuro.neuromatch.io/tutorials/W1D3_GeneralizedLinearModels/student/W1D3_Intro.xhtml
- https://dartbrains.org/content/GLM.html
- [don’t watch it unless you are very comfortable with GLM] https://www.youtube.com/watch?v=NFeGW5IjUoI&list=PL9YzmV9joj3FIkQwVcfj1VsLV_pj6Cwcr
- https://open.win.ox.ac.uk/pages/fslcourse/website/online_materials.html - the fMRI 1 & 2 sections here provide a good overview on how GLMs and fMRI data analysis

Retinotopic Maps:

- https://www.youtube.com/watch?v=MhFJIgeY-ZY [a must watch!]
- https://www.youtube.com/watch?v=HCLgh9AwvrQ [good visualization]

Hemodynamic response function:

- https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4254893/
- https://andysbrainbook.readthedocs.io/en/latest/fMRI_Short_Course/Statistics/03_Stats_HRF_Overview.html

W1D2 Project Diary:

1. We decided on using the HCP retinotopy dataset.
2. We think of comparing different ML algorithms and modeling in general (i.e., our primary focus is on comparing different methods to approach one problem) to maximize our learning experience.
3. For tomorrow, our objective is for everyone to look at the dataset (HCP retinopathy) Guide to choosing an FMRI dataset — Neuromatch Academy and come up with a problem and a hypothesis, and in general, just explore the data.

W1D3 Project Diary:

Ideas:

1. CNNs - how much of the input signal (convolutional layers) do we need to estimate the pRF accurately? https://www.youtube.com/watch?v=70A3uYfM1qA
2. GANs - could be a bit complicated, but a really neat way to use generative models and unsupervised learning https://www.youtube.com/watch?v=TpMIssR
3. # Compute median spike count
4. median_spike_count = np.median(total_spikes_per_neuron) # Hint: Try the function np.median
5.
6. # Visualize median, mean, and histogram
7. with plt.xkcd():
8. plt.hist(total_spikes_per_neuron, bins=50, histtype="stepfilled")
9. plt.axvline(median_spike_count, color="limegreen", label="Median neuron")
10. plt.axvline(mean_spike_count, color="orange", label="Mean neuron")
11. plt.xlabel("Total spikes per neuron")
12. plt.ylabel("Number odhco
13. Maybe according to the time series, we can predict not only the type of stimulus but also its direction! expanding to what degree or at which position while rotating (e.g., at 6,9 or 12 degrees in a clock) !. (but the problem is maybe algorithms will use other brain areas but the visual cortex to predict the direction of the movement, but that's maybe for discussion later on). Note: The stimulus data is provided by NMA.
14. Compare moving angles to static angles and see if the activity we see comes from movement or the angle itself
15. The main paper can be used to see what kind of analysis they used, and we can also look at their freely available code. Example of a previous analysis of the same set of data GitHub - mszinte/HCP_dataset_analysis: Analysis codes for the HCP 7T retinotopy task (not by the authors, we still need to look for that one).
16. How can we incorporate the analysis of the pRF size and eccentricity?
17. In the vision there is this thing where input from the left side of the visual field crosses onto the right side of the brain and vice versa. Maybe we could do something about cross-site interactions - like modeling/making an NN from the dataset, and then running them on test data with only half of the stimulus? To see if, and what influence there is (binocular vision is super interesting, but maybe this example is not very good for the project).

Summary of the day:

1. We decided to start with analyzing the data in a simple manner and then gradually increase the complexity (so we start with GLM, supervised learning, then move all the way to neural networks and deep learning).
2. We did some brainstorming on different ideas to expand our project furthermore (check the first section).

3. **Plan for tomorrow:**
    a. We will start applying the GLM already (after attending today's tutorial). And we will make a common collab. (Pick the right GLM according to the data type and the proper regularization method).
    b. General brainstorming of different research questions and hypotheses.

W1D4 Project Diary:

The Goal of the project:

Our learning goals for this project are to understand machine learning techniques and implement neural networks. Thus, instead of looking for an original question in the field of vision that can be answered by our set of data, and due to a lack of experience in this field, we decided that our goal is to discuss the models themselves and why they are more or less accurate. We will put a hypothesis before using each model of what we think the model's results will be And why? According to the type of data and the amount of it. Which will help us understand the models more and enhance our learning experience.

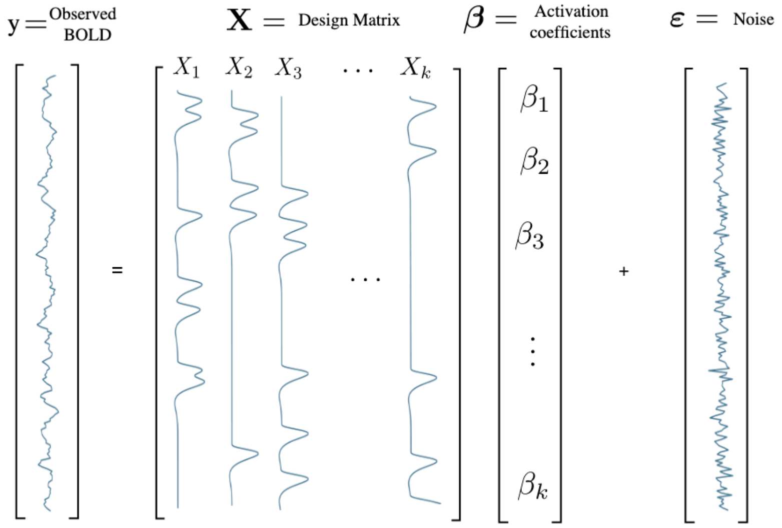
For this purpose, we will need to read more about each model we use and predict its outcomes.

How to approach the project:

1. Start with descriptive models: GLM and dimensionality reduction.
2. Then move to the fancier models, e.g., ANNs, CNNs, and GANs.
3. An important question we should consider is how to visualize our highly dimensional data?

And the Problem we encountered while looking at the data:

Please watch the GLM in fMRI youtube video linked above to gain an understanding of how GLM is used in fMRI analysis!



The Stimuli used during the tasks are continuous and multi-dimensional, i.e., they can't be represented as discrete events that can be used for further prediction purposes and convolving the stimulus tract with the HRF to make a predicted signal.

Also, how are we going to encode the stimuli? For all stimuli, is it informative to use the height and width as x parameters in the design matrix? Or, for the expanding circle stimulus for instance, would it be more informative to calculate the area of the outer and inner rings?

Deconvolving - we could try to deconvolve the neuronal data with the HRF to get approximated discrete stimulus events - moments when the brain area responded to whatever was happening in the stimulus. It would make an easier output for a model that we would train, but we can skip this by just adding convolution to the model.

Next steps:

1. Tomorrow we will apply GLM and PCA algorithms.
2. Tomorrow as well, we are contacting our mentor and TAs to have a meeting on Monday at 11:00 am GMT+1 for 30-60 mins.
3. Prepare before Monday:
    a. Results from GLM and PCA.
    b. General conceptualization of the neural networks (especially after WD5 deep learning tutorials). For this, we need to read some reviews and write some ideas.
    c. three slides presentation, showing intro to retinotopic mapping, our goal, results from the descriptive models, hypothesizing why we will use the different neural networks and any problems we are encountering now.

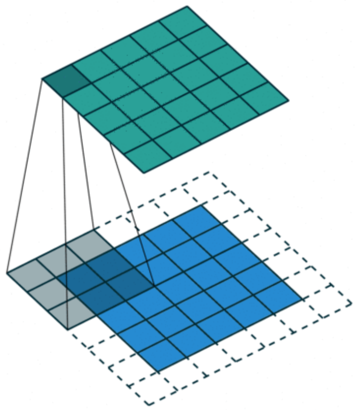
W1D5 Project Diary:

Summary of why we want to use CNNs? And what we are going to do with it? (Normative modeling, encoding, decoding)! rls83@cam.ac.uk

A convolutional neural network (CNN) is a class of artificial neural networks (ANN) typically applied to analyze visual imagery. Briefly, they work by progressively combining small portions of the input into features and then combining those features into larger and larger features as the layer progresses (watch this video for a complete overview https://www.youtube.com/watch?v=UzfRnZqPbCI). Note: General ANNs have no knowledge about the overall structure of your data. CNN's are designed for data that are spatially related (i.e., pictures, where each pixel is related to the pixel around it).

CNN components:

- Input layer/ blue matrix - tensor with shape (number of inputs) x (input height) x (input width) x (input channels)
- Filter/kernel/gray matrix - designed to detect a specific type of feature in the input (e.g., center-surround, Gabor)
- Output layer/green matrix - activation/feature map with shape (number of inputs) x (feature map height) x (feature map width) x (feature map channels)



Note the difference between decoding and encoding tasks:

- Decoding = model predicts visual stimuli based on brain responses (neural activity, could be synthetic or empirical)
- Encoding = model predicts brain responses based on presented visual stimuli

In W1D5 tutorial 3, we trained a deep network to discriminate stimulus orientation, then compared the DNN neurons' representations of the stimuli to the brain's representation of the stimuli (as measured empirically by fMRI; an encoding task). In the scope of our project, we are planning to train and optimize DNNs to perform similar tasks and expand on this:

1. Object recognition: Train a DNN to discriminate the 4 different shapes present in the HCP retinotopy stimulus dataset (ring, wedge, diagonal bar, horizontal bar).
2. Naturalistic task: Train a DNN to discriminate stimulus movement (i.e., expanding vs contracting ring; CCW vs CW wedge; horizontal vs vertical vs diagonal movement of the bar)
3. Train network to directly estimate stimulus area (exp/contracting ring); angle of stimulus (CCW/CW)

For each of these DNNs:

1. How do the network unit representations compare to brain representations as measured by fMRI?
2. How can we explore different neural network architectures to optimize correlation with empirical BOLD activity (i.e., optimal depth & width of layers)?

*****Note**: Training on different tasks could lead to different representations of the oriented grating stimuli, which might match the observed V1 representations better or worse**

General Ideas we discussed:

- We will start with only predicting the type of stimulus (whether it appears or not, all or none), and compare the different types of stimuli (bars moving, circles expanding, and wedges rotating).
- We can also divide the project into encoding and decoding parts.
- We can also divide the project into the prediction of movement, change in the area, rotation, orientation, and so on.

Tasks to be done:

Task	Assigned to
Code for the contracting circle.	Khansa - done
Code for the rotating wedges.	Ray - done!
Code for the moving bars.	Pawel - done
Make a presentation on Google slides:	Khansa - done
Read more on GANs (a review maybe), and generally come up with a hypothesis	Everyone!
Summarize why CNNs, and classic ANNs (also try to make a slide for that, so we are finished with this)!	Ray - done!

W2D1 Project Diary:

Summary: we have a representation and basic GLM for each category of stimulus

Meeting with Peter Vavra

- Check stimulus representation with what is in the literature
- Need tuning curves - i.e. what voxels "prefer" what feature
    - Eccentricity of ring
    - Phase (angle) of wedge
    - Similar logic for moving bar
- Use these tuning curves (B-maps or T-maps) as starting point for DNN
    - Searchlight sizes & shapes
    - Which voxels and how many do we need in our DNN layers (spatial aspect)
    - Intuition - more voxels in ML algorithm, the more chances the algorithm has to accurately predict. BUT losing spatial specificity. By tweaking what voxels we allow the algorithm to be trained on, we're gaining the possibility to make claims about how spatially localized information is
    - Whole-brain decoding analyses = not limiting voxels, gain info about different brain areas
    - Choice of the volume of voxels limits you to certain aspects of data the algorithms can pick up
- Can group stimuli based on features of interest
    - Interesting to explore different aspects of features (is it about the shape, or the animate vs inanimate)
    - For our learning purposes - pick 1 (the standard way of doing retinotopy) and delve deep into that
    - What we start with now is what we will have to work with in the future
- Ideas (time tradeoffs)
    - Pick multiple algorithms (LR, LDA, SVMs…) and one searchlight size
    - Pick 1-2 algorithms and multiple searchlight sizes
    - Decoding mindset - can we tell what the visual stimulus was
    - Can confidence ratings be read out of encoding models? - look for literature for what is currently out there (conceptually)
    - Compare layers of DNN to neural representations from different areas of the brain

Step 1: Find a phenomenon and a question to ask about it

- What aspect of the data needs modeling?
    - Technicalities of ML & phenomenon of model fitting - how can we develop the best ML model to represent visual processing in the human brain as it has been empirically shown in the literature

  - How do different algorithms allow us to take simple inputs & infer higher-order features from them?
- Define an evaluation method - How will we know our modeling is good?
  - ML evaluation techniques: accuracy, sensitivity, specificity
  - Compare representation in model layers to neural activity in HCP dataset
- Experiment to test model - create stimulus combining features (ring, wedge, bar) & see how the model performs

```
**Make sure to avoid the pitfalls!**

Question is too general

Precise aspect of phenomenon you want to model is unclear

You have already chosen a toolkit

You don't have a clear goal

You don't have a potential experiment in mind
```

Ideas to execute	**Later Maybe**
Decoding (GLM, SVM, CNN)	Encoding (Bayesian model)
Whole brain	specific areas
Naturalistic task	Object recognition
Position of object (spatial)	Movement of object (temporal)
CNN spatial	CNN temporal (sliding window)
Conventional CNN	spline-based convolutional neural network (SplineCNN)

Algorithms to use:

1. Classic DNNs
2. CNN's (sliding window to incorporate temporal dimension? https://bmcbiomedeng.biomedcentral.com/articles/10.1186/s42490-019-0003-2)

Step 2: understanding the state of the art & background (lit review)

- Survey the literature
  - What's known?

    - What has already been done?
    - Previous models as a starting point?
    - What hypotheses have been emitted in the field?
    - alternative/complementary modeling approaches?
- What skill sets are required?

Step 3: determining the basic ingredients

- What parameters/variables are needed?
    - Constants
    - Do they change over space, time, condition…?
    - What details can be omitted?
    - constraints/initial conditions
    - Model inputs/outputs

- Look for standardized versions of stimulus expressions
- Sliding window CNN - do we want to try to later include the temporal aspect in the model?

- Variables needed to describe the process to be modeled?
    - What can be observed/measured? Latent variables?
    - Where do these variables come from?
    - Do any abstract concepts need to be instantiated as variables?

- Ways of measuring model/ANN performance - correlations, accuracy, what else?

```
**Make sure to avoid the pitfalls!**

I'm experienced, I don't need to think about ingredients anymore

I can't think of any ingredients

I have all inputs and outputs

can't think of any links (= mechanisms)
```

Step 4: Formulating specific, mathematically defined hypotheses

- Think about the hypothesis in words by relating ingredients identified in step 3
    - What is the model mechanism expected to do
    - How are different parameters expected to influence model results?
- Express the hypothesis in mathematical terms by giving the ingredients variable names
- *more precise hypotheses = easier to justify model!

- Hard to describe with our knowledge right now, let's come back after doing lit review

Tasks for tomorrow (Literature Review):

Task **(please summarize your findings below)**	**Who can do it**
Different ML algorithms (SVM, GAN,..) **suitable if you are just starting with us**	Done
How the stimuli are represented in the literature	Done
B-maps or T-maps	
Sliding window CNN	
Ways of measuring model/ANN performance, correlations, accuracy, what else?	

W2D2 Project Diary:

Papers and short summaries:

- We can use SVM as a supervised learning algorithm, it has already been proven to work within the setting of retinotopy: check this old paper Inverse retinotopy: Inferring the visual content of images from brain activation patterns - ScienceDirect whats very promising as well is that they are using rotating wedges and contracting and expanding circles as well in their first experiment.
- We can also use Geometric Deep learning which is an algorithm used to predict the retinotopic organization in this recent neuroimage paper Predicting the retinotopic organization of human visual cortex from anatomy using geometric deep learning - ScienceDirect.spline-based convolutional neural network (SplineCNN) (apparently it's a special type of CNN)!, (Pytorch Geometry).
- One last thing we can do incase we have more time and after we have the tutorial on Bayesian modling, is to use the hierarchical bayesian models for prediction (might n´be from encoding only): Evaluation of hierarchical Bayesian method through retinotopic brain activities reconstruction from fMRI and MEG signals - ScienceDirect and Bayesian analysis of retinotopic maps | eLife (elifesciences.org).
- For now I think our focus should be on the following in respect to the models we will use: mainly three models for decosing: GLM, SVM, and CNN. and if we have time encoding by a bayesian model.

Future Plan:

1. Work on only three models together: GLM, SVM, and conventional CNN and compare results.
2. For GLM: we will use our own coding of the stimuli for the seek of time constraints, we will code the same stimulus in several ways and compare results.
3. CNN spatial then sliding window.

For tomorrow:

1. Improve wedge stimulus coding then do CNN on that.
2. Start SVM implementation.

W2D3 Project Diary:

What we did today:

1. CNN tuning curves
2. Wedge stimulus coding finished
3. SVM implemented

Later on:

1. Refine hrf.

For tomorrow:

1. Implement everything (do GLM, SVM, and CNN) on all the stimuli.
2. List problems of implementation.

W2D4 Project Diary:

What we did:

1. Visualized the tuning curves in heat maps form.
2. Applied SVM on the counter clockwise wedges.
3. Used PCA to reduce the dimensions of the voxels as a feature.

Ideas:

- Visualize the heat map in a 3D brain animation.
- Use first classical neural networks, because we dont know how to do CNNs on time serieses. Check this out when we start CNN: How to Develop Convolutional Neural Network Models for Time Series Forecasting (machinelearningmastery.com)
- Or we can use CNN for encoding and SVM for decoding
- Lasso regularizer!

For tomorrow:

1. SVM on all the stimuli
2. Classic neural network

W2D5 Project Diary:

What we did today:

1. Implemented GLM on all the stimuli but CCW.
2. Implemented CNN normative encoding model
3. Implemented SVM for both object recognition, and different features of the stimuli.
4. HRF fixed the negative part.
5. Met peter.

For Monday:

1. Abstract.
2. Implement Peters suggestions.
3. Finish everything we started.,

Summary of Peter's suggestions:

GLM: we could calculate the "preferred stimulus feature" for each voxel, by running GLM with specific features as parameters, e.g. 45 degree angle, 90 degree angle, to get tuning curves for voxels, and then measure the specificity of each voxel, to get voxels that care the most about stimulus features, and then use those to feed other models.
(https://pages.ucsd.edu/~msereno/papers/HumanRetin95ms.pdf)

PCA: right now we are doing PCA over space, getting the characteristic time series for each principal component; we could do PCA over time to get the characteristic spatial response for each principal component, basically showing us which areas care about what we are looking at.

For other types of dimensionality reduction, we could use lasso regularization or e.g. 20% of most caring voxels from the preferred stimulus map or the voxels from PCA characteristic spatial responses.

CNN: we can fit CNNs to do what we want in 3 main ways:

- Make a model with several outputs, each relating to a stimulus feature (even if it's not present in all the cases)
- Make a model with one output, and train it on all kinds of data (but it'll create twisted representations in the hidden layers describing all the features at once)
- Make one model for each feature and combine them

For other CNN things ask Ray, I didn't actually note most of the stuff here lol

SVM: for the step where we distinguish stimuli types, we need to account for lag related to HRF. We can also compare doing PCAs for each stimulus separately to doing it for all of them together, to see if using global patterns is better than more localized ones specific to each stimulus (probably better to make it task specific)

Since GLM doesn't really fit in our project so far, we could either try using those preferred stimulus maps, or if time allows, include encoding models in our project (stimulus -> fMRI data). We can ignore a lot of discretization/lag problems by including HRF convolution in those models, basically making them predict the "stimulus events", which when convolved with HRF give our raw fMRI data. Having both would be cool for a GAN model, but dunno if we have time.

W3D1 Project Diary:

Abstracts

1:

Retinotopic mapping involves measuring the spatial organization of the human visual cortex using functional magnetic resonance imaging (fMRI) while subjects view differentially spaced and moving stimuli. The Human Connectome Project collected a freely available dataset of 181 healthy subjects measuring retinotopic mapping using 7T fMRI. In this project, we analyze these data using generalized linear modeling (GLM) and supervised learning to investigate how the brain represents visual information. We used GLM to identify brain regions that are most responsive to specific features of the visual stimuli. We then trained support vector machines (SVMs) to identify visual stimuli features based on neural activity. Finally, we trained deep neural networks (DNNs), including a convolutional neural network (CNN) layer, to recognize visual stimuli, then compared the stimulus representations of the 'neurons' in this model to actual neural activity as measured by fMRI. Together, these results increase our understanding of the spatial representation of visual stimuli in the cortex, and give insight into how the brain computes such information.

2:

Computational neuroscience provides a wide range of tools for modeling the mechanisms of neural action. Deep neural networks in particular come in many varieties and may seem applicable to almost anything. Given the breadth of choice, it might be, however, unclear which model to utilize, especially when one is a beginner in the field. Here we provide a comparison of several models, namely the generalized linear model (GLM), support vector machines (SVM), and convolutional neural networks (CNN), performing a simple visual stimulus decoding task. Using neural data from 7T fMRI retinotopic mapping taken from Human Connectome Project (HCP), we train the models to predict one of six different stimuli, including rings changing sizes, rotating wedges and moving bars. In order to introduce a layer of complexity to the task, we are also trying to decode information about the features of the stimulus, such as angle or area, in parallel to stimulus discrimination. Then we compare the accuracy and the ease of implementation between the models. We expect to see CNN as the best performing model, as its structure is most fitted to processing visual data. Those results would let us suggest that different models have different areas where they thrive, which can be understood through experience with using those models. We also hope this project will let us increase our personal understanding of those as well.

Final:

In computational neuroscience, identifying the precise model that can answer specific questions is key to identifying neural mechanisms.Given the breadth of tools available, it might be, however, unclear which model to utilize, especially when one is a beginner in the field. To explore which aspects of neural data different models can explain, we inspect the Human Connectome Project (HCP) 7T retinotopic mapping data set. The dataset used functional magnetic resonance imaging (fMRI) to measure the spatial organization of the human visual

cortex while subjects view differently spaced and moving stimuli. We investigate the utility of several models, namely the generalized linear model (GLM), support vector machines (SVM), and deep neural networks (DNN), to perform visual stimulus decoding and encoding tasks. Specifically, we used GLM to identify brain regions that are most responsive to specific features of the visual stimuli. We then trained SVMs to identify visual stimuli features based on neural activity in a decoding task. Finally, in a simple encoding task, we trained DNNs, including a convolutional neural network (CNN) layer, to recognize visual stimuli, then compared the stimulus representations of the 'neurons' in this model to actual neural activity as measured by fMRI. Together, these models explain different aspects of visual stimuli representation in the brain, and give insight into how the brain computes such information.

How can we use models to explore visual stimulus representation in the brain?

How do different models explain different aspects of neural data?

W3D2 Project Diary & W3D3 Project Diary & W3D4 Project Diary:

Sections	Tasks	Done or not
HRF	Fix the negative part	Done - stopped working
	Get the voxel specific HRF into GLMs	Done - stopped working
Stimuli	Code CCW	Done
	Code bar2	
GLM	Calculate tuning curves for feature for each voxel!	Done
	Get specificity of the tuning curves	Not enough time
SVM	PCA for each of the stimuli separately	Done
	Visualize all the SVMs [python - Visualize 2D / 3D decision surface in SVM scikit-learn - Stack Overflow]	I think its difficult to visualize our SVMs, because we dont want to classify, but rather predict just like a simple regression problem. So maybe we can just print the report in table and the confusion matrix. Note SVMs are used for both classification and regression
	Account for HRF lag	
	Combine stimulus guessing SVM with each of the small ones	
CNN	Make the encoding CNN (stimulus ↠ fMRI data)	

Other	Make the presentation	
	Organize the common coding file	In progress