**A. What are the phenomena?**

*The BOLD response of the brain to natural visual stimuli is predicted (or characterized) by the features learnt based on an artificial neural network representation*

**B. What is the key scientific question?**

*Is there a strong biologically plausible relationship between a convnet(Alexnet/ResNet-18) and the human visual pathway? Does each layer in the neural network have a close analogue that maps to the visual pathway?*

**C. What was our hypothesis?**

*Earlier visual pathway layers (i.e., V1 to V3/4) are easier to predict with earlier ANN layers, however, higher order processing of visual stimuli (i.e., MST to FFA) will be harder to predict with the last layers of the network.*

**D. How did your modeling work?**  
*We used a pretrained resnet18 model to extract features from visual stimuli, storing the activations for each of the model’s forward pass layers. Given the large size of these output vectors, we ran dimensionality reduction using PCA, first identifying the n-components that would explain 95%+ of the variance. Following these, L2 linear regression models we run on each layer to predict the BOLD response and compare to each ROI voxel activations from human participants.*

**E. What did you find? Did the modeling work?**

**F. What can you conclude?**

**G. What are the limitations and future directions?**

*It will be interesting in future research to see the performance of video classification neural networks like 3D resnet in predicting BOLD response to the different videos. In addition we believe that this kind of neural networks might also be able to deal with the BOLD delay in the response*

**Abstract Draft 1**

Artificial neural networks (ANN) have been shown to predict (or characterise) BOLD response activations in the brain following visual stimuli, through a process of feature extraction that mimics the human visual system. For example, the architecture of some deep learning models have been found to be good artificial representations of the structural connections in the ventral visual pathway (v.v.p.) However, knowledge representation in the visual system integrates with various cognitive systems (e.g., attention, working memory and semantic memory) to process information received. As such, models trained on still images are unable to capture the more dynamical activations that the human visual system processes in reality. Here we wanted to investigate if there is a strong biologically plausible correlation between activations from a residual CNN (Resnet18), and the v.v.p.. In particular, we wanted to know if each Resnet layer can show a close analogue that maps to the areas of the v.v.p. in response to more dynamical natural stimuli. We hypothesise that while early CNN layers map well with early cortical areas of the v.v.p.; higher order processing of visual stimuli will not be predicted as accurately in later CNN layers as these layers are fed by many systems, some of which could not be taken into account in the current model. We used a pretrained Resnet18 model to extract features from short video clips (provided by the Algonauts Challenge 2021) and stored the output activations for each of the model’s main layers. Dimensions of output were reduced through PCA prior to voxel activation prediction with L2 regression models. We then compared the predicted fmri responses of each layer with voxel activations in each region of interest using a representational dissimilarity matrix. This study is important for advancing our integration of biological systems with artificial networks, which are cornerstones to real world applications such as cortical prosthetic devices. Future studies could investigate the performance of video classification neural networks like 3D resnet in predicting BOLD response to the different videos.