Passive Attention in Artificial Neural Networks Predicts Human Visual Selectivity

Thomas A. Langlois,* H. Charles Zhao,* Erin Grant, Ishita Dasgupta
Thomas L. Griffiths† & Nori Jacoby†

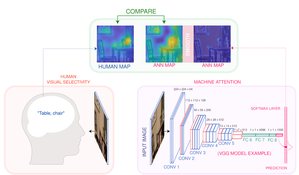
*T.A.L. and H.C.Z. contributed equally to this work. †T.L.G. and N.J. contributed equally to this work.

NEURAL INFORMATION PROCESSING SYSTEMS

BACKGROUND

- ***We compare human visual selectivity to neural network attention***

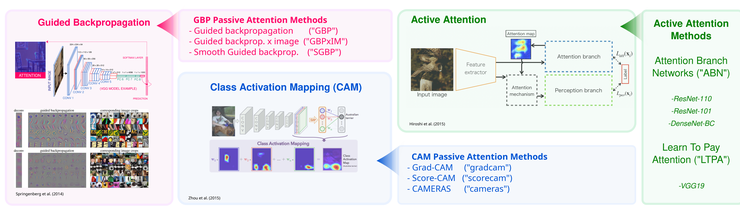
HUMAN VISUAL SELECTIVITY & ANN ATTENTION



NEURAL NETWORK MODELS

Model	# params
CIFAR AlexNet	2.50M
CIFAR VGG-19 (w/ BatchNorm)	20.09M
CIFAR ResNet-110	1.73M
ImageNet AlexNet	61.10M
ImageNet VGG-16 (w/ BatchNorm)	138.37M
ImageNet ResNet-101	44.55M
ImageNet EfficientNet-B0	5.29M
ImageNet ViT-S/16	22.05M
Places AlexNet	58.50M
Places ResNet-50	24.26M
CIFAR LTPA VGG	19.99M
CIFAR ABN ResNet-110	3.06M
CIFAR ABN DenseNet-BC ($L = 100$, $k = 12$)	1.12M
ImageNet ABN ResNet-101	62.58M

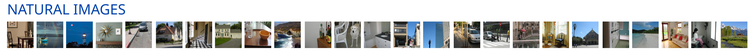
GUIDED BACKPROPAGATION (GBP), CLASS ACTIVATION MAPPING (CAM), ACTIVE ATTENTION



EXPERIMENTS

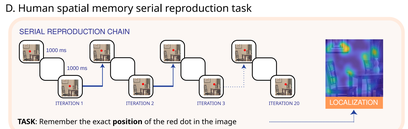
- ***We obtained human visual selectivity estimates for 6 behavioral tasks***

NATURAL IMAGES

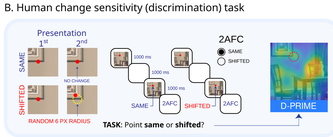


A. Human image patch recognizability ratings task

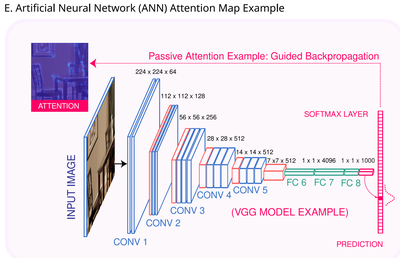
D. Human spatial memory serial reproduction task



B. Human change sensitivity (discrimination) task



E. Artificial Neural Network (ANN) Attention Map Example



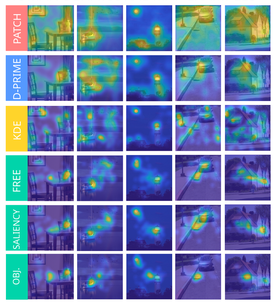
C. Human eye-tracking: free view, object & saliency search

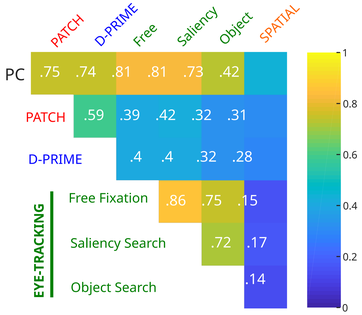
Koehler et al. (2014)

HUMAN BEHAVIORAL MEASURES

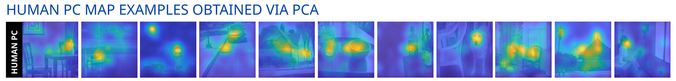
- ***Human visual selectivity estimates show systematic variation***

HUMAN BEHAVIORAL MAP EXAMPLES





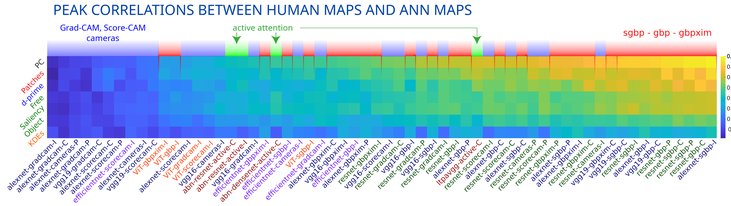
HUMAN PC MAP EXAMPLES OBTAINED VIA PCA



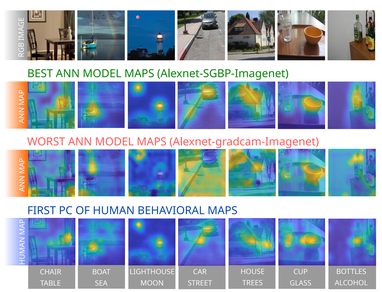
HUMAN AND NEURAL NETWORK ATTENTION

- ***GBP attention predicts shared component of human measures***

PEAK CORRELATIONS BETWEEN HUMAN MAPS AND ANN MAPS

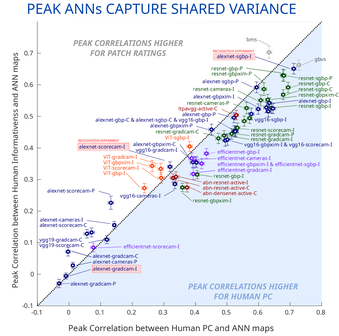


EXAMPLE ANN MAPS AND HUMAN PC MAPS



Human labels for original images

PEAK ANNs CAPTURE SHARED VARIANCE



HUMAN RECOGNITION EXPERIMENT

- ***Neural network attention influences human recognition performance***

HUMAN RECOGNITION: MASKING

EXPERIMENT DESIGN

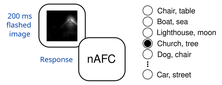
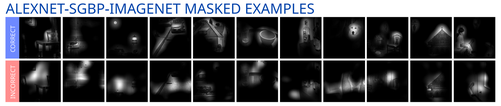
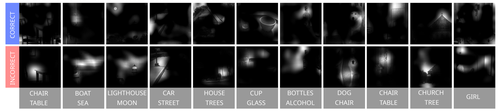


IMAGE EXAMPLES

ALEXNET-SGBP-IMAGENET MASKED EXAMPLES

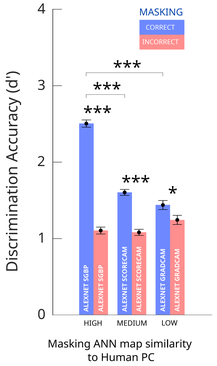


ALEXNET-GRADCAM-IMAGENET EXAMPLES



Human labels for original images

RECOGNITION RESULTS

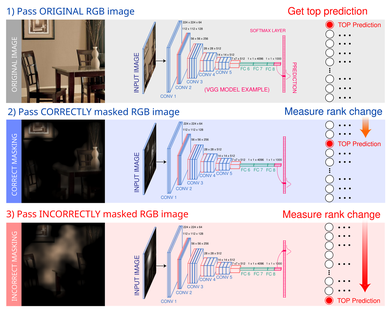
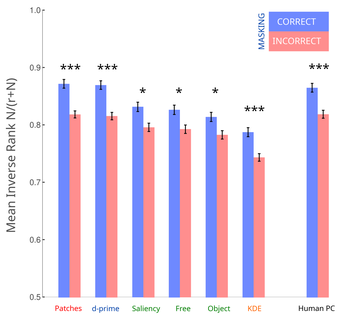


NEURAL NETWORK RECOGNITION EXPERIMENT

- ***Human visual selectivity directly influences neural network recognition***

Inverse Rank: $N / (r + N)$ (NUMBER OF CLASSES; TOP PREDICTION RANK)

MASKING AND RANK CHANGE



MEAN INVERSE RANK RESULTS