Behavior of neural network models performing visual search tasks

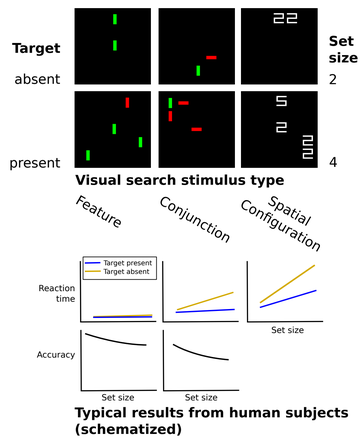
David Nicholson, Astrid Prinz. Emory University, Biology department.

nicholdav@gmail.com nickledave nicholdav

Introduction

Visual search

- **Lab visual search tasks** focus on how visual system integrates **single features** (color, orientation)



- Real-world visual search involves **object recognition**

Is there a plane? (easy)

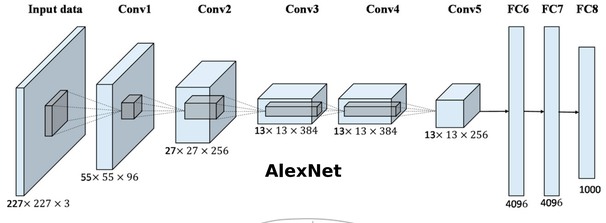
Is there a garbage can? (difficult)

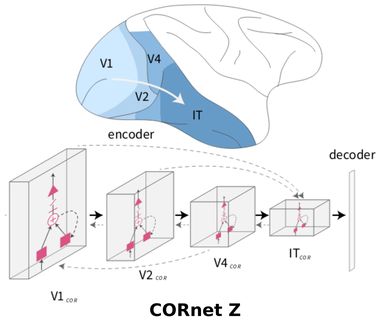
Images from Pascal VOC dataset, used to create the Visual Search Difficulty dataset.

Question: can **visual search** behavior be explained by **object recognition**?

Currently, best models of object recognition are neural networks

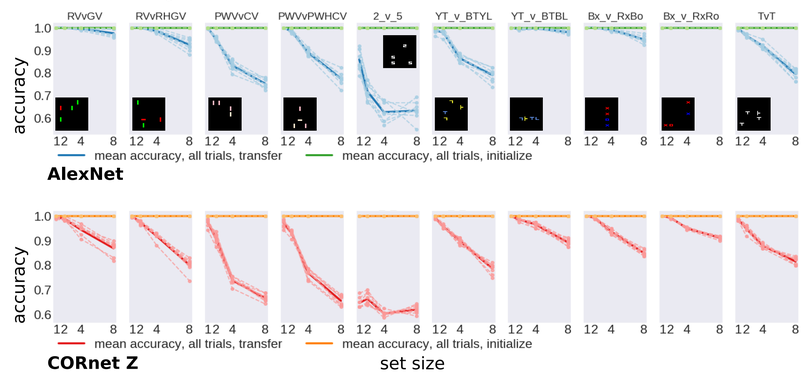
- fit by optimizing for image classifcation
- predict both brain activity + behavior





Results

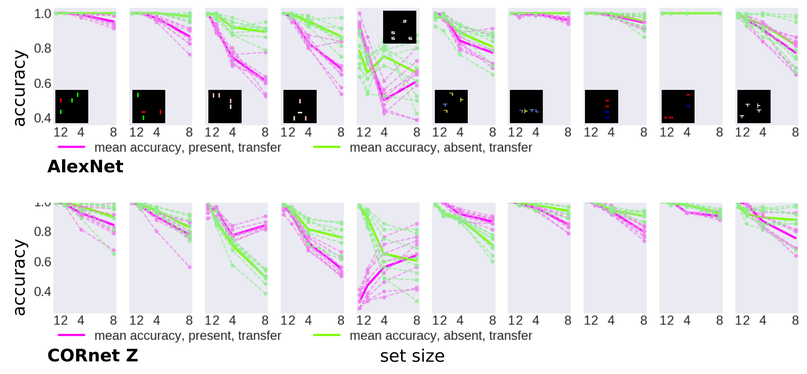
Object recognition models predict accuracy on visual search task



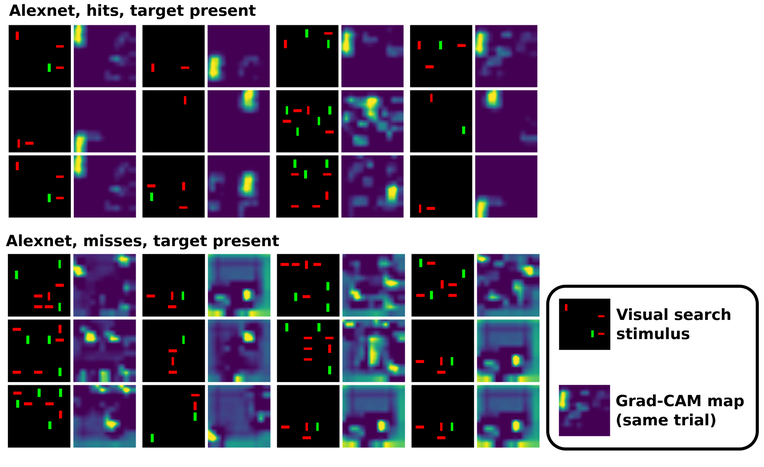
Further analysis suggests this is because of how neural networks map pixels to output

Accuracy always lower when target is present

Note that network sees many more distractors during training

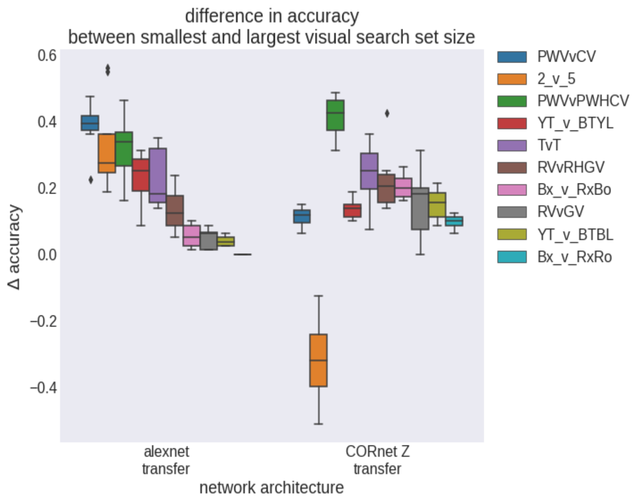


On miss trials, target pixels not important for output



Results, continued

Object recognition models similarly rank impact of target-distractor pairs on accuracy



Conclusions

- Biases learned from natural images cause models to "behave" like humans performing visual search
- Closer analysis and visual explanation techniques suggest neural network models map pixels to output, do not represent target as humans might do
- In spite of this, these models may predict target-distractor discriminability in humans
  - similar to ideal observer models

Methods

Adapt object recognition models, i.e. neural networks pre-trained on image classifaction, to visual search tasks, using transfer learning methods from machine learning. For details, please see: https://github.com/NickleDave/visual-search-nets

Acknowledgements

Research funded by the Lifelong Learning Machines program, DARPA/ Microsystems Technology Office, DARPA cooperative agreement HR0011-18-2-0019. David Nicholson was partially supported by the 2017 William K. and Katherine W. Estes Fund to F. Pestilli, R. Goldstone and L. Smith, Indiana University Bloomington