

# A Bayesian Model of Visual Attention

Sharat Chikkerur, Thomas Serre, Cheston Tan & Tomaso Poggio  
Center for Biological and Computational Learning, MIT

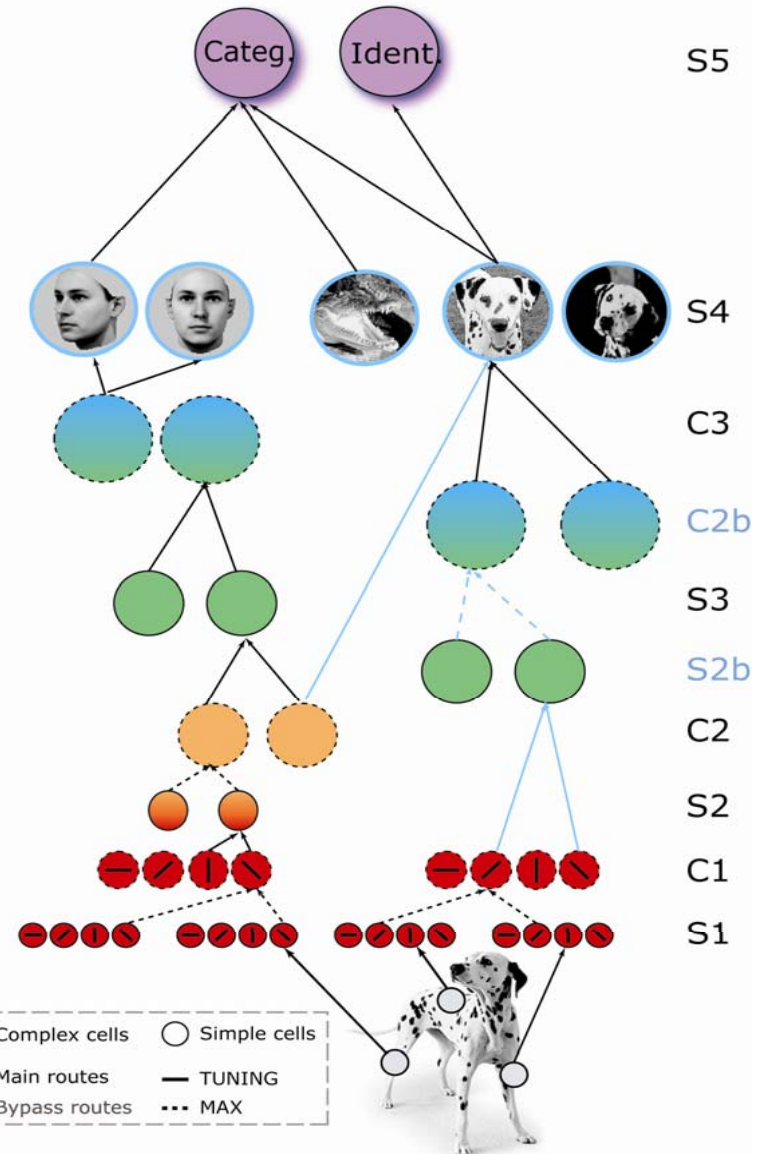
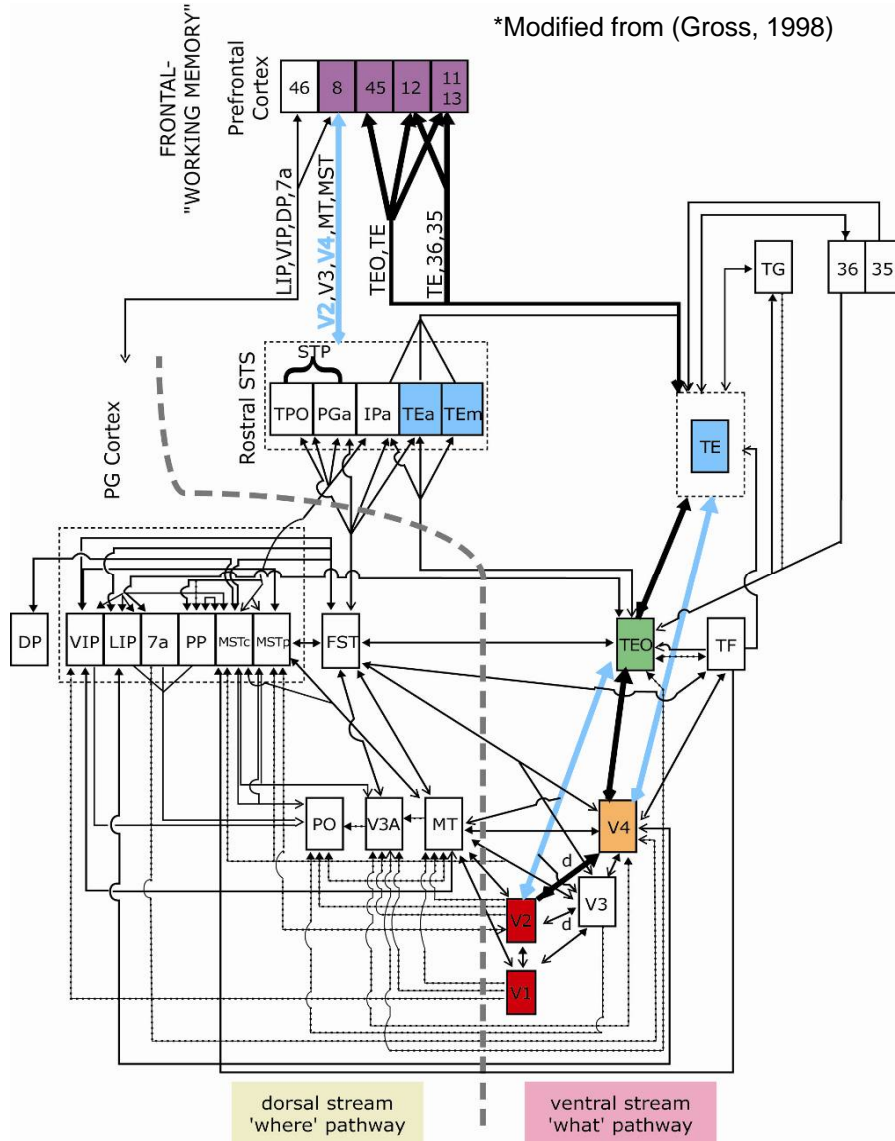
# Outline

- Introduction
  - Limitations of feed-forward processing
  - Role of attention
- A computational model of attention
- Applications
  - Modeling human eye-movements
  - Object recognition under clutter

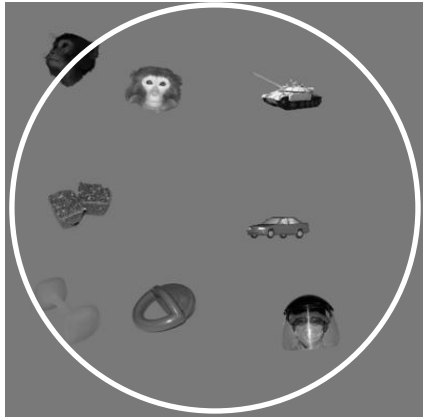
# Outline

- Introduction
  - Limitations of feed-forward processing
  - Role of attention
- A computational model of attention
- Applications
  - Modeling human eye-movements
  - Object recognition under clutter

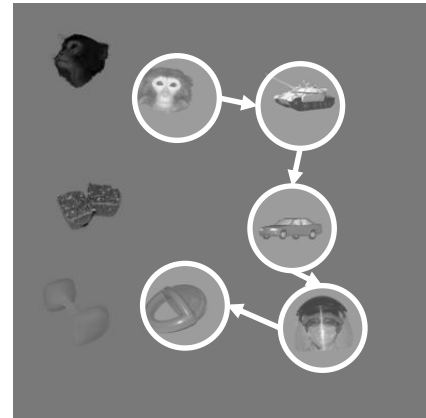
# Feed-forward processing



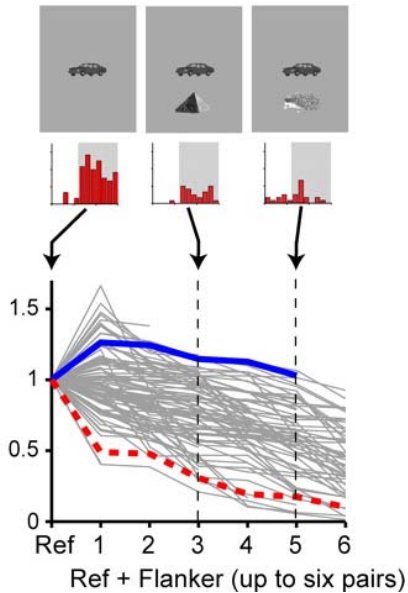
# Role of attention



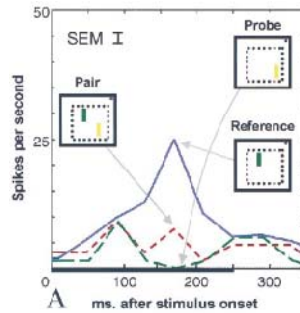
Parallel processing (No attention)



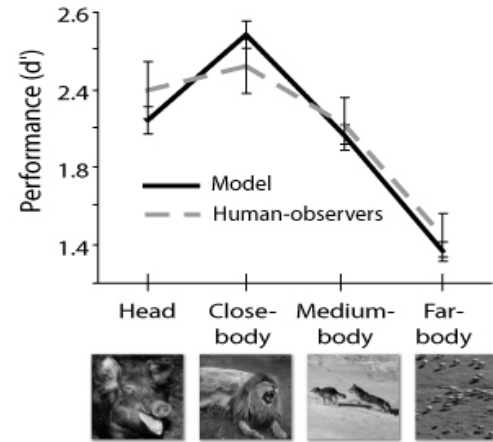
Serial processing (With attention)



Zoccolan Kouh Poggio DiCarlo 2007

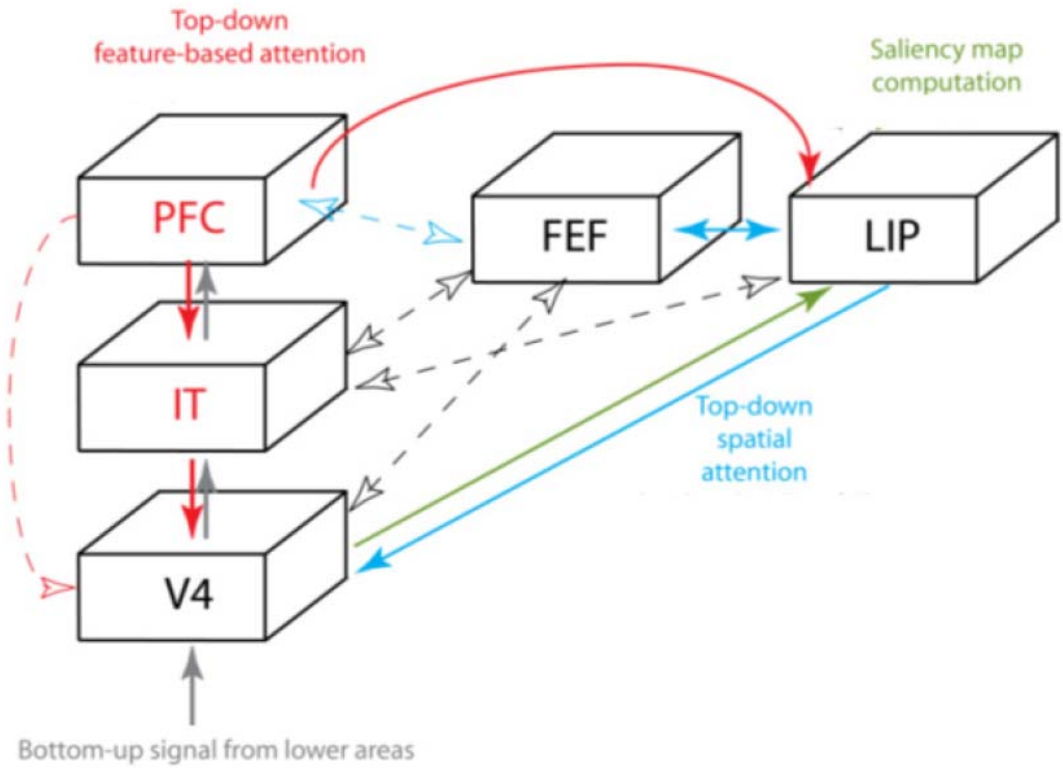
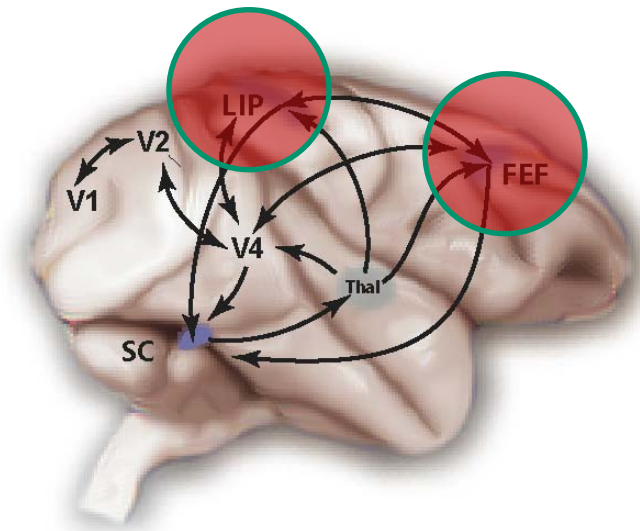


Reynolds Chelazzi & Desimone 1999

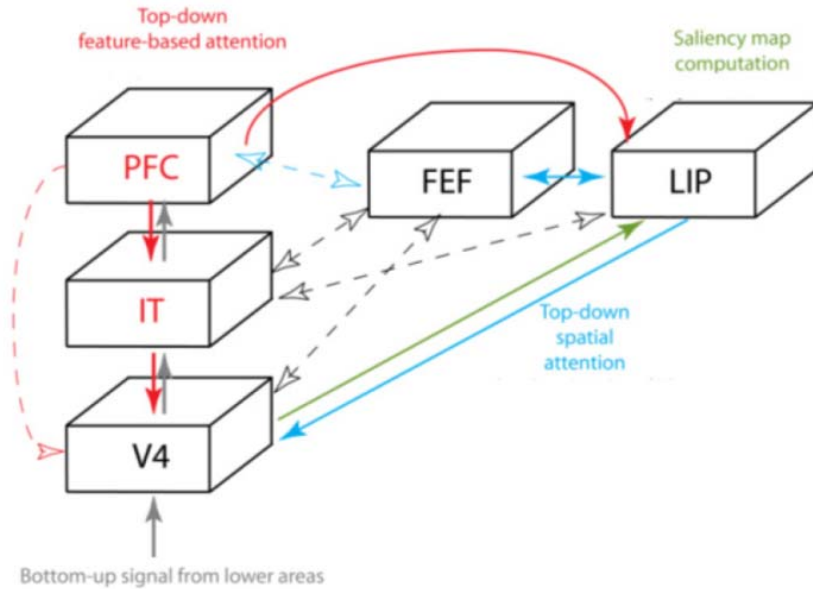


Serre Oliva Poggio 2007

# Biology of attention



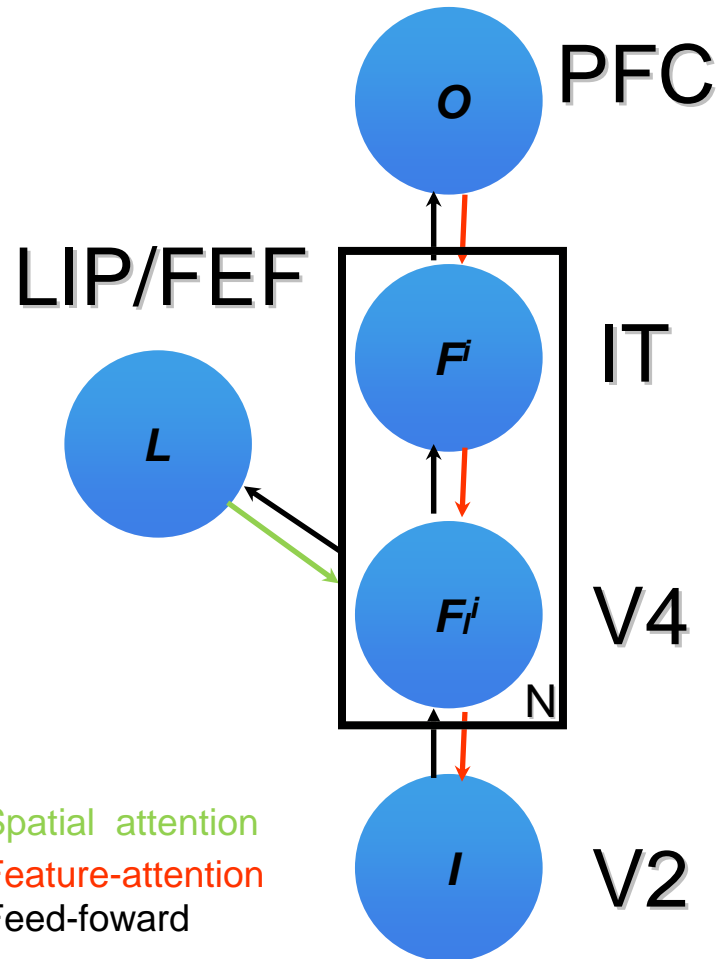
# Attention as Bayesian inference



- We use a Bayesian framework to model the interaction between the ventral stream and LIP/FEF

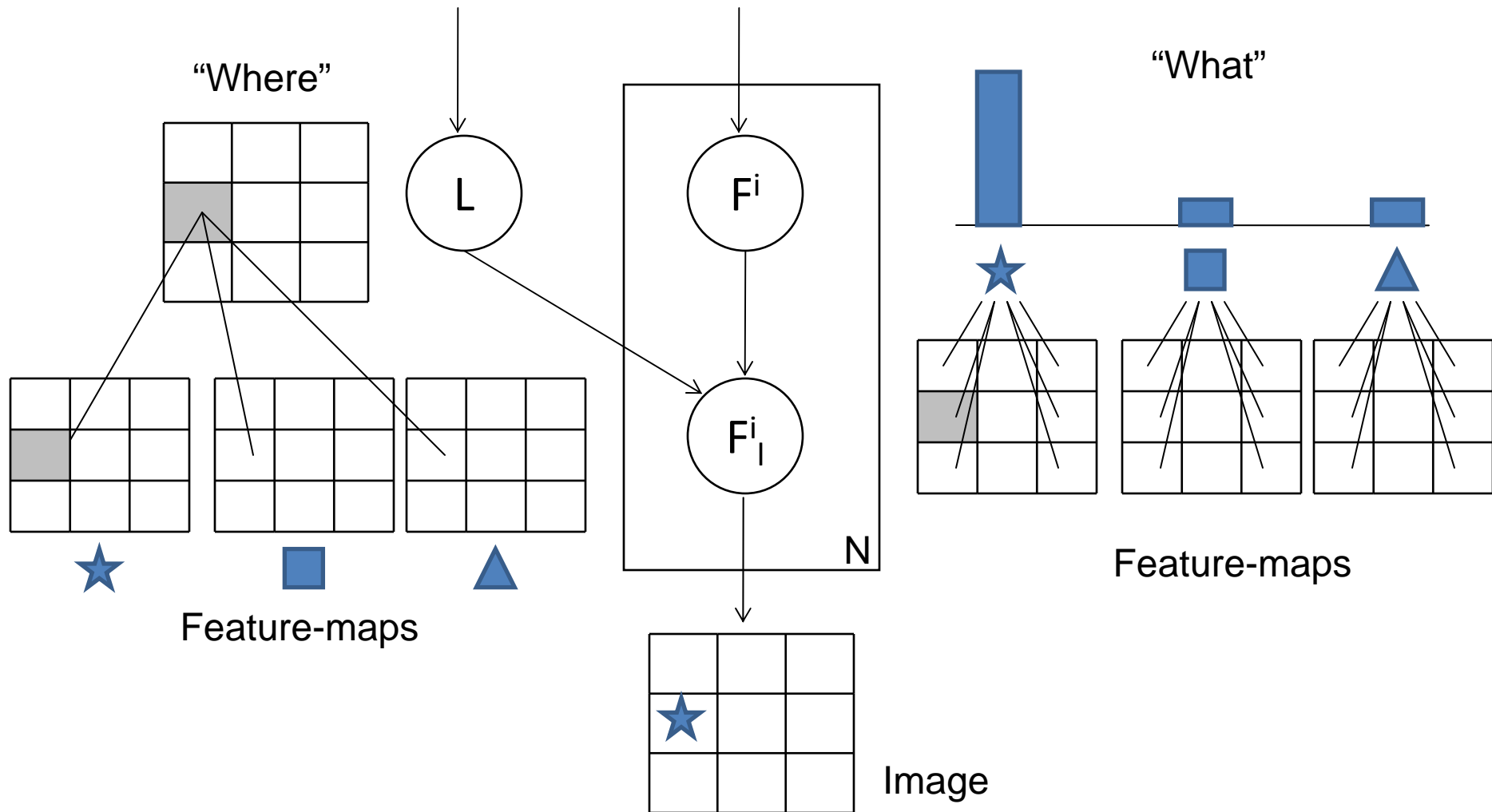
- Feed **forward** connections within the ventral stream are modeled as **bottom-up** evidence. **Feedback** connections from higher areas are modeled as **top-down** priors.

- The posterior probability of location generates a task-based saliency map.



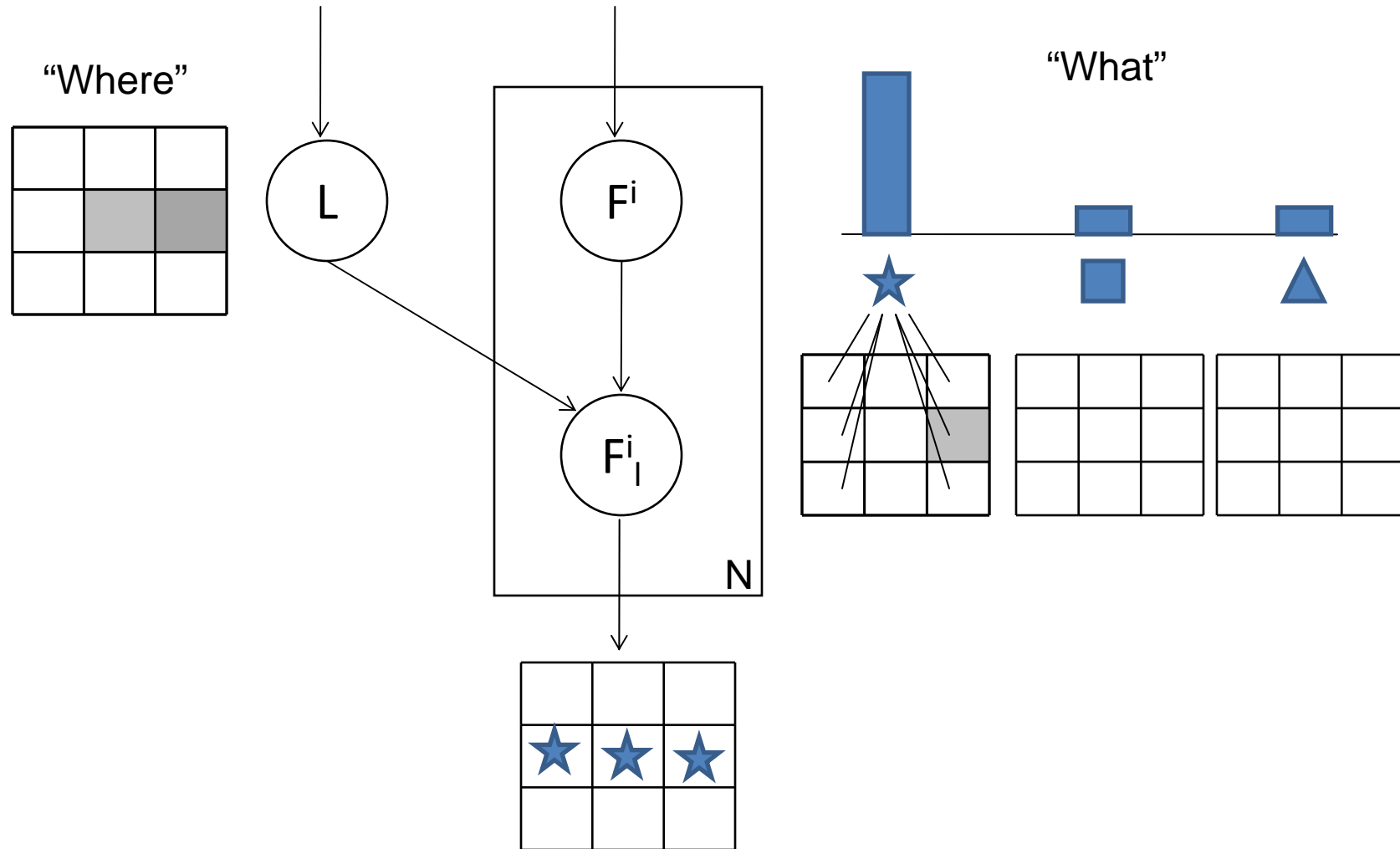
■ Spatial attention  
 ■ Feature-attention  
 ■ Feed-forward

# Model description

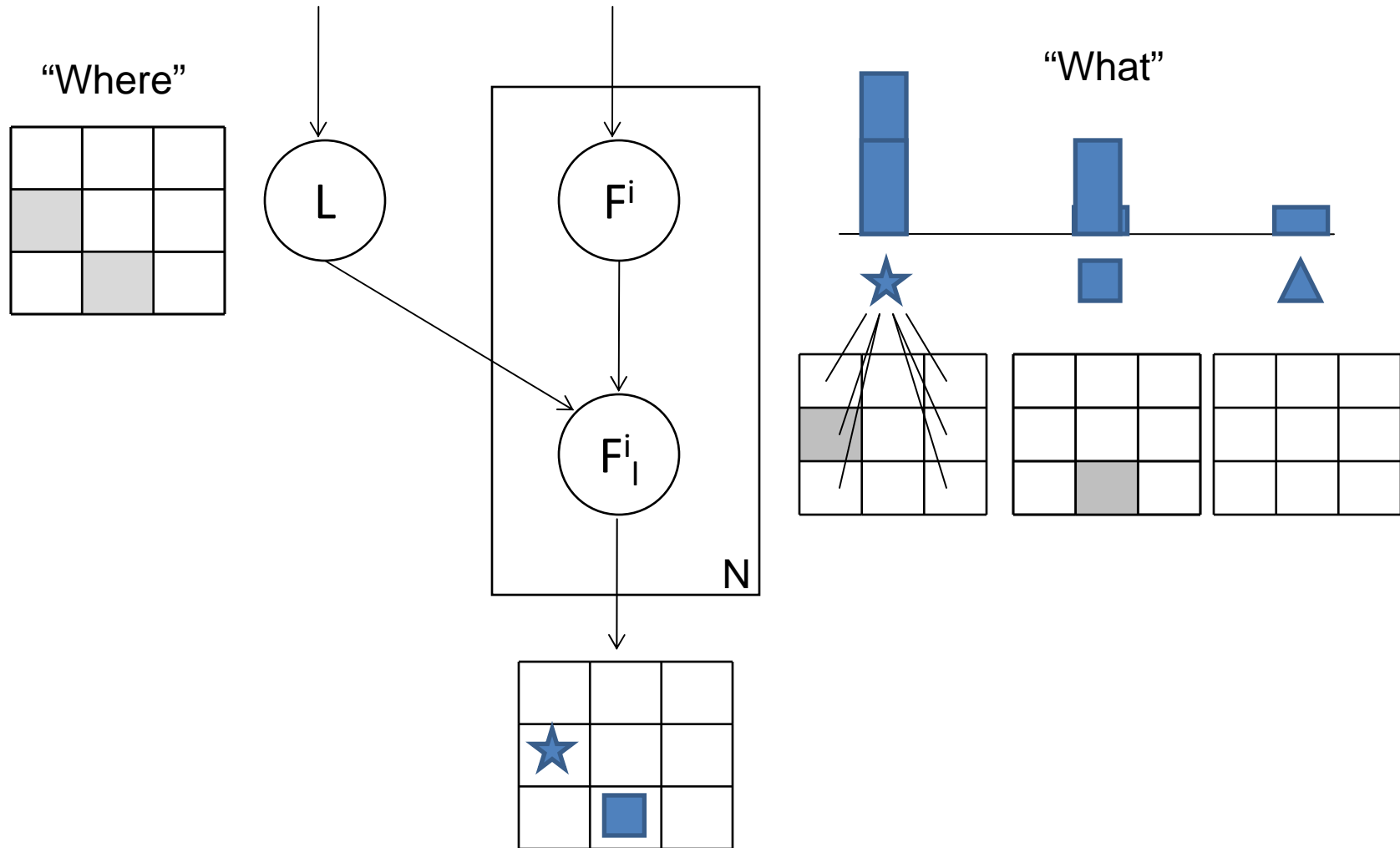




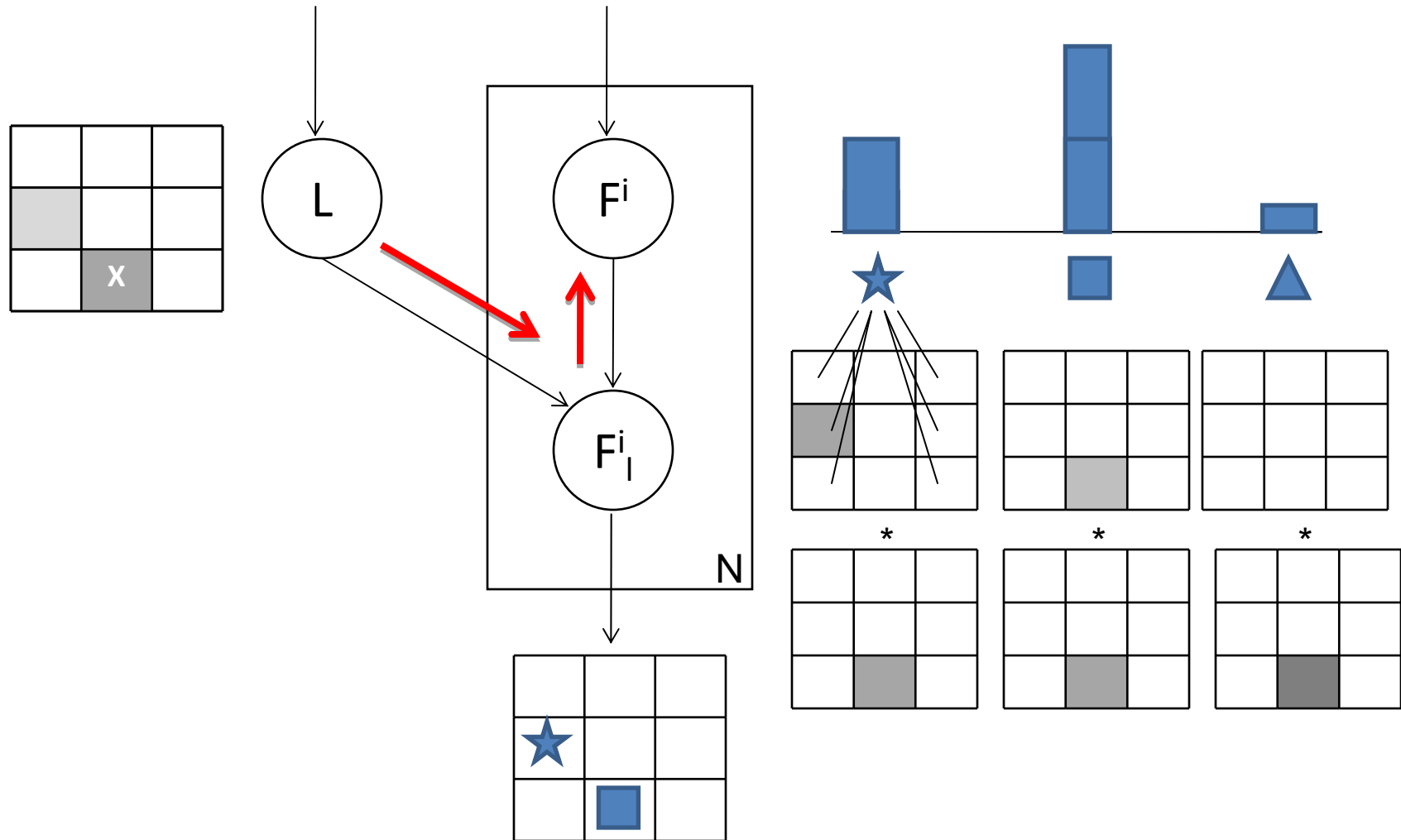
# Model properties: invariance



# Model properties: crowding

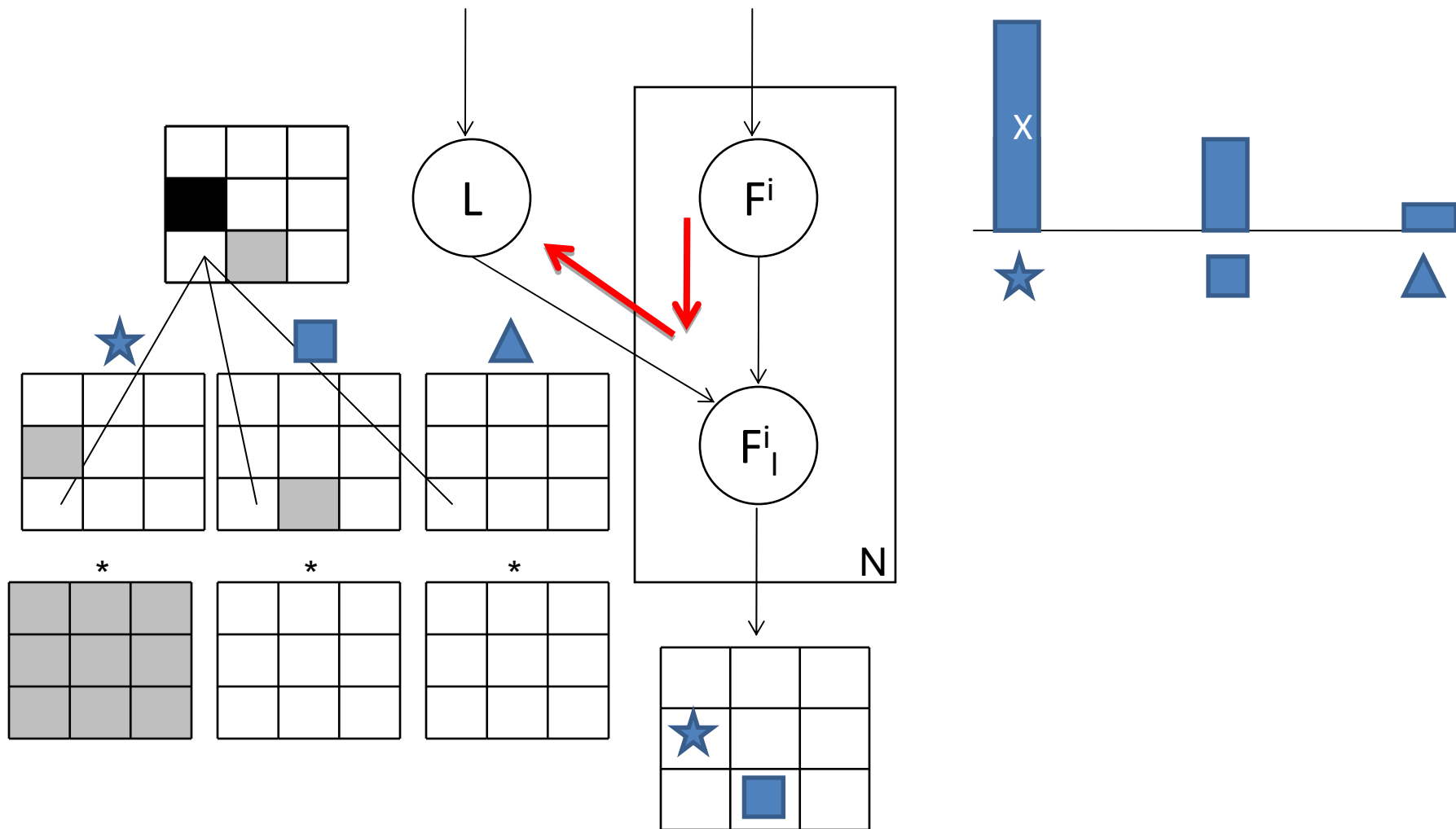


# Model: spatial attention



■ **What** is at location X?

# Model: feature-based attention



▪ **Where** is object X?

# Outline

- Introduction
  - Limitations of feed-forward processing
  - Role of attention
- A computational model of attention
- **Applications**
  - Modeling human eye-movements
  - Object recognition under clutter

# Matching human eye-movements



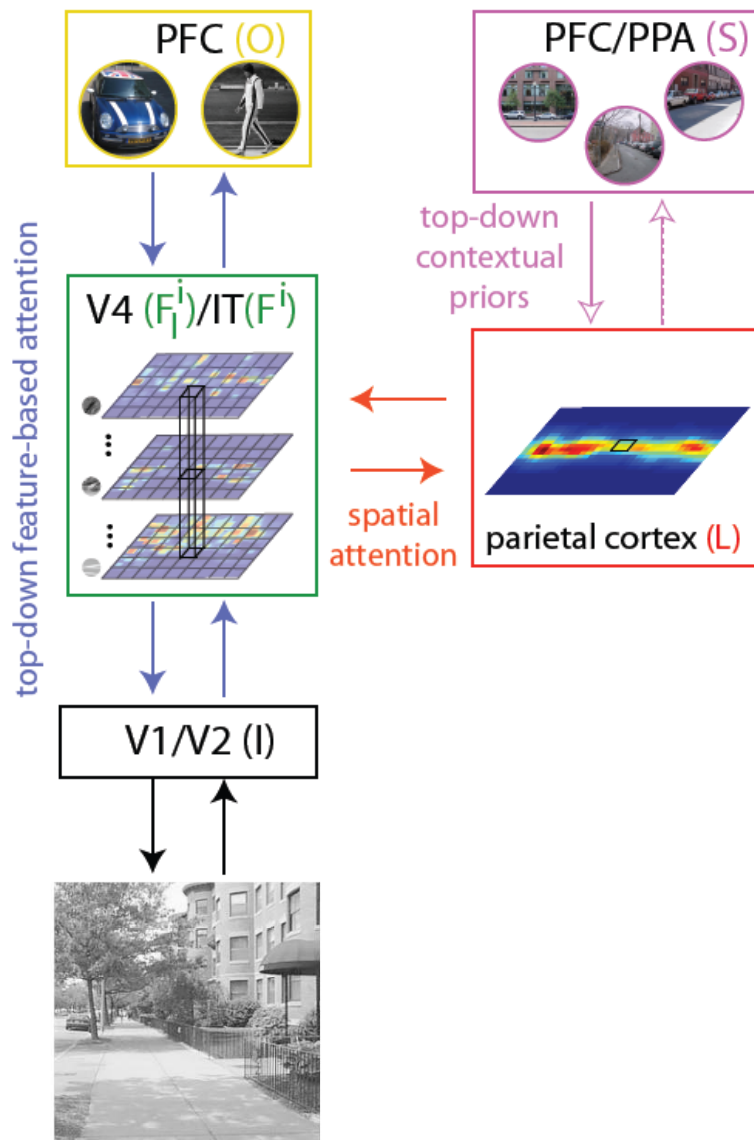
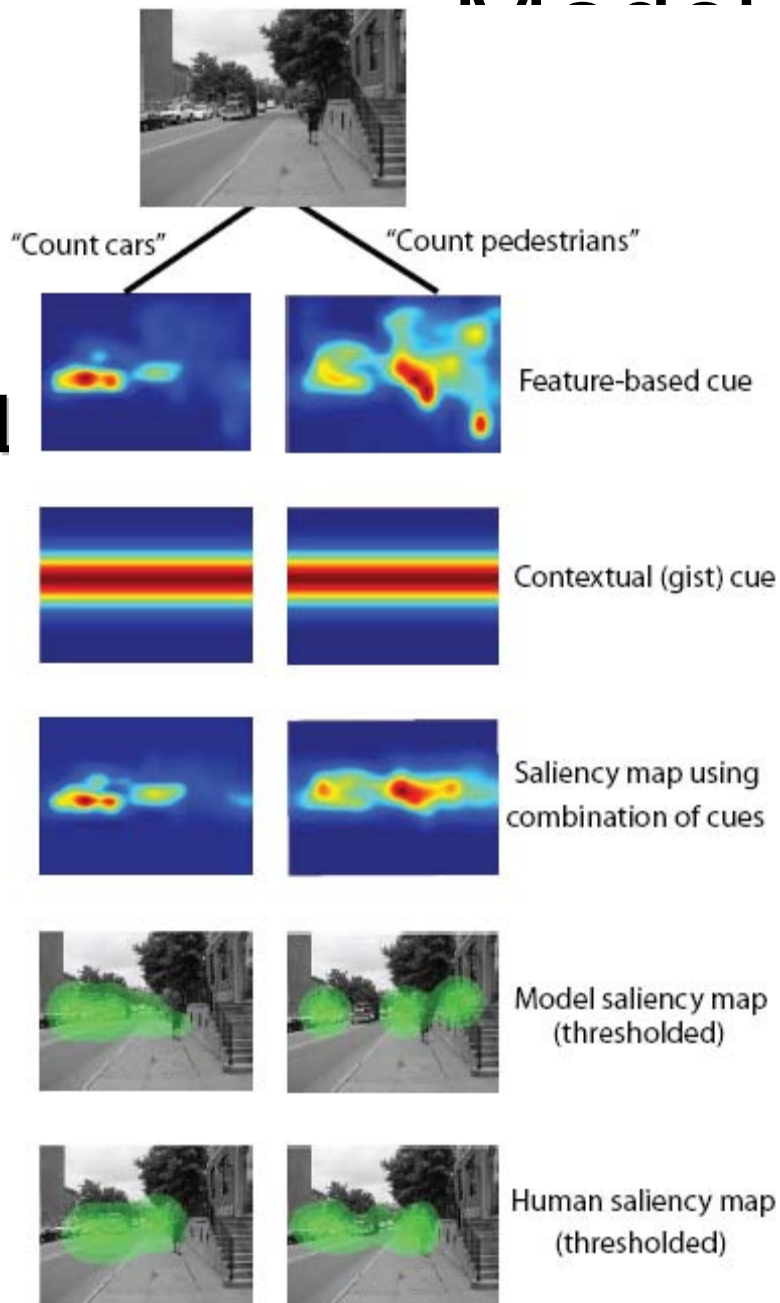
Car search



Pedestrian search

- **Dataset**
  - 100 CBCL street-scenes images having cars & pedestrians
  - 20 images with neither objects
- **Experiment**
  - 8 subjects were shown these 120 images in random order.
  - Each image in the stimuli-set was presented twice
  - The subjects were asked to count the number of cars/pedestrians
  - For each of these block trials, the subject's eye movements were recorded using an infra-red eye tracker.

# instantiation



# Examples

Car Search



Car Search



Car Search



Pedestrian Search



Pedestrian Search



Pedestrian Search



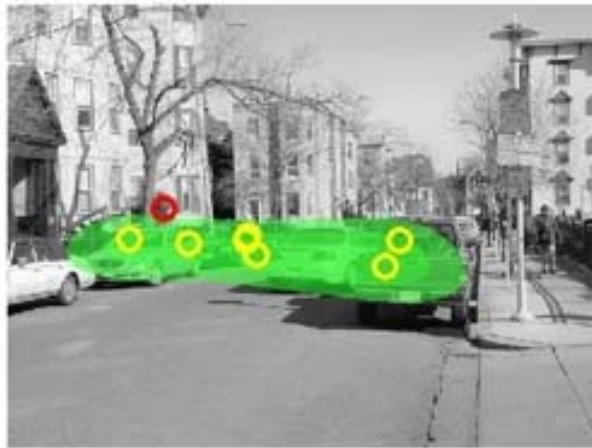


# Examples

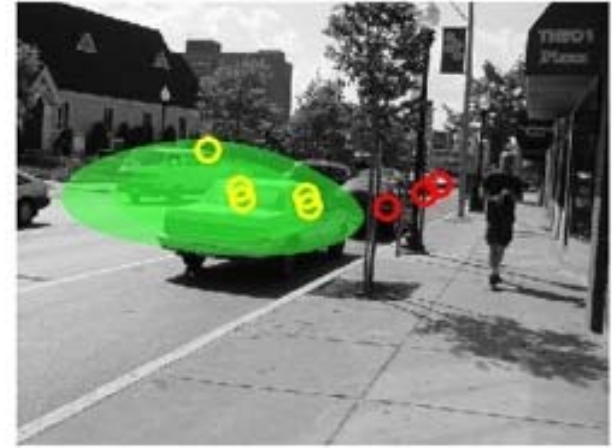
Car Search



Car Search



Car Search



Pedestrian Search



Pedestrian Search

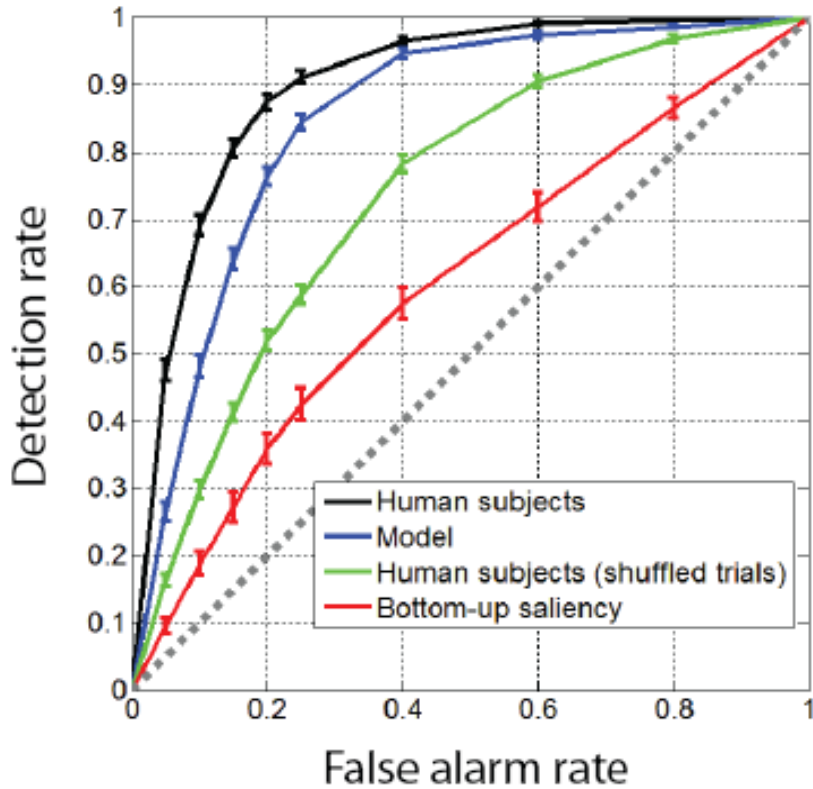


Pedestrian Search

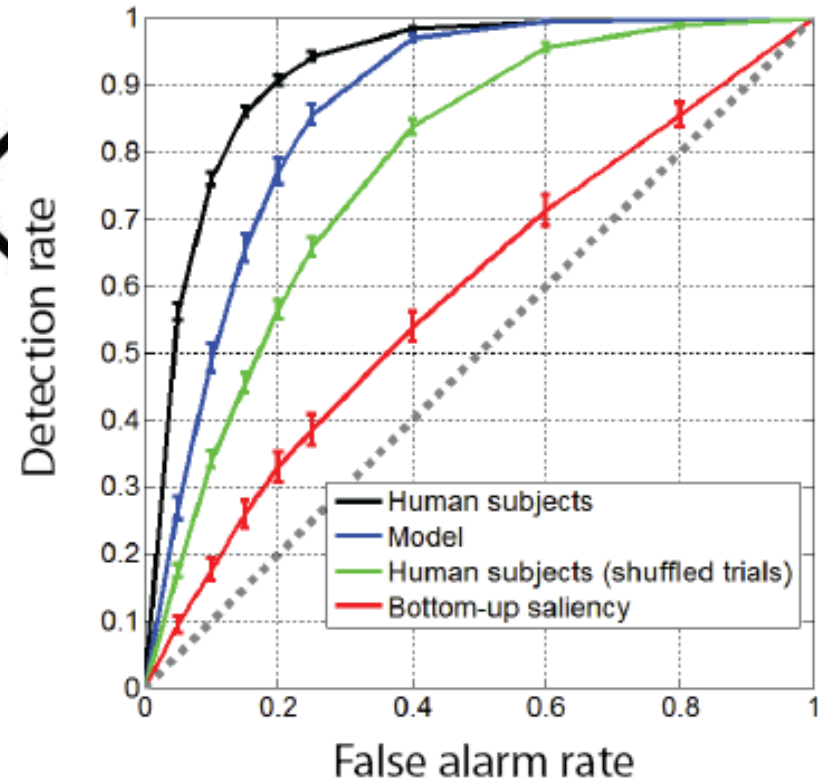


# Quantitative evaluation: ROC

## Car Task

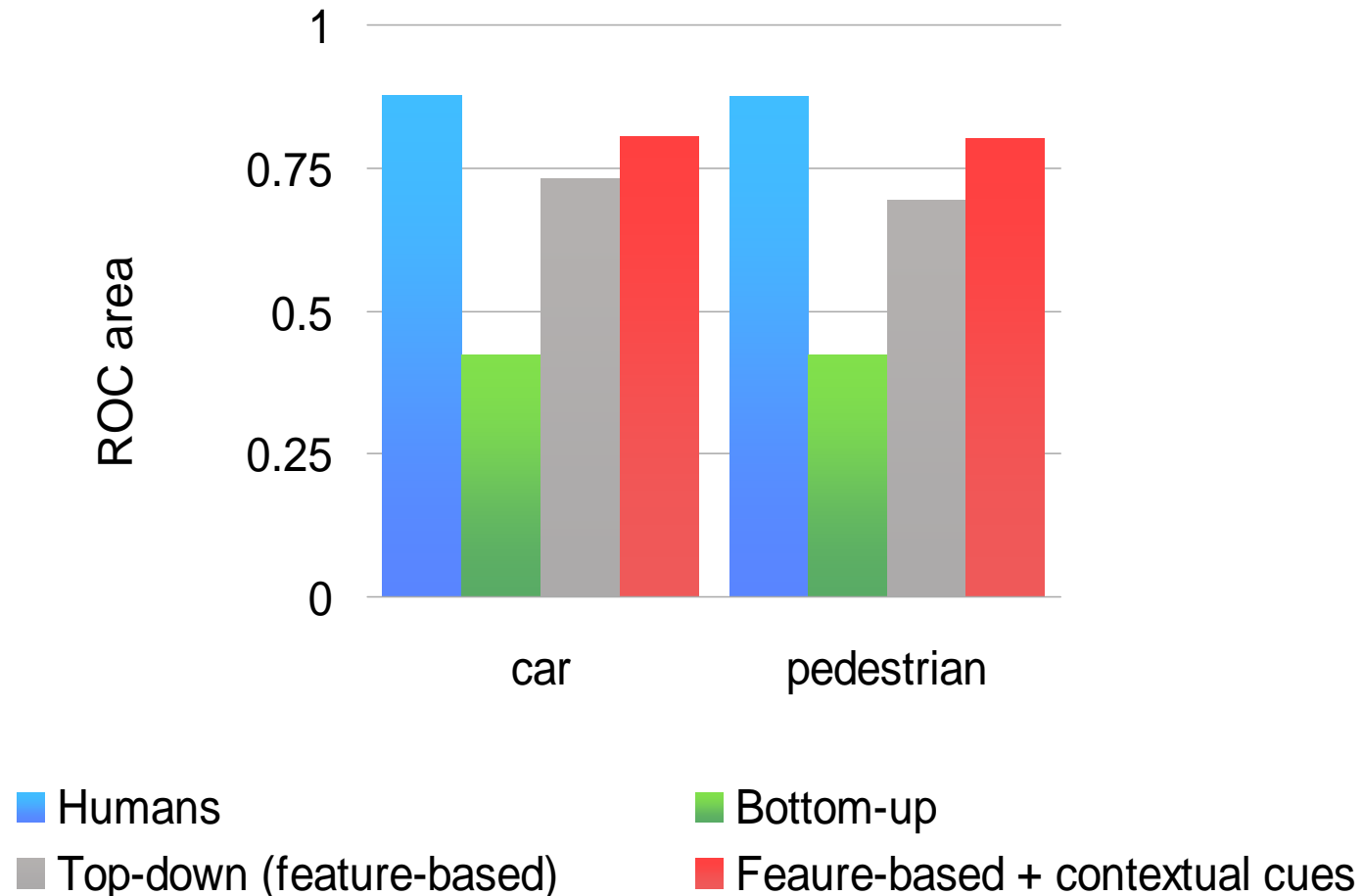


## Pedestrian Task



# Quantitative evaluation: ROC

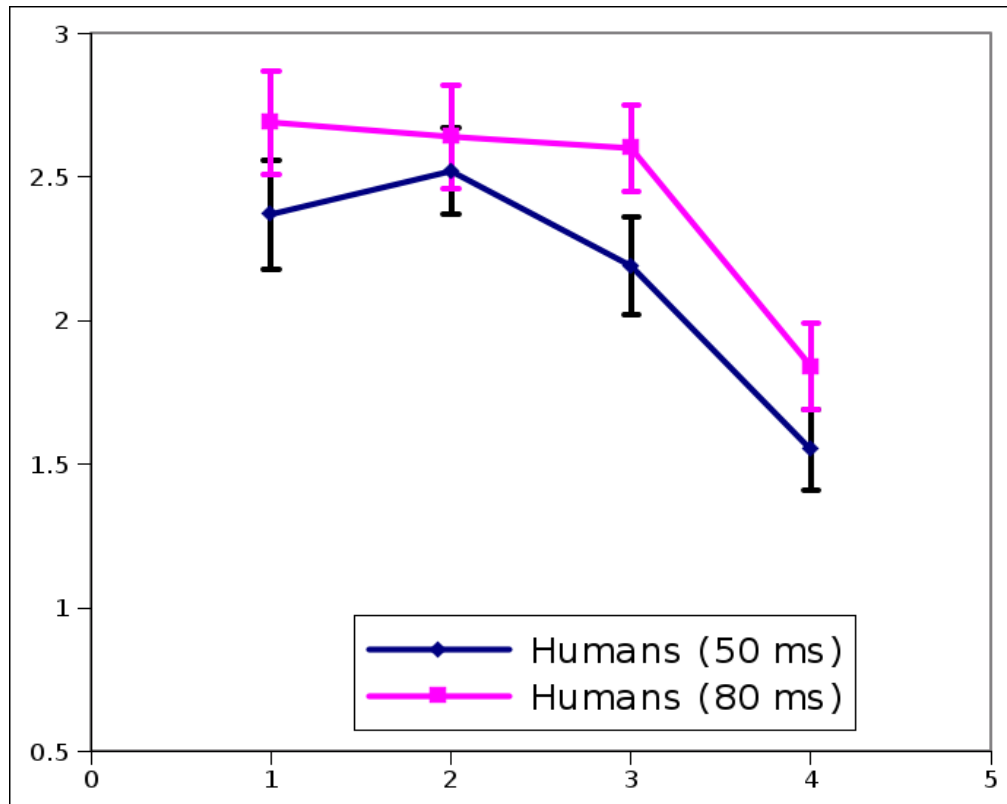
Integrating (local) feature-based + (global) context-based cues accounts for **92%** of inter-subject agreement!



# Outline

- Introduction
  - Limitations of feed-forward processing
  - Role of attention
- A computational model of attention
- **Applications**
  - Modeling human eye-movements
  - **Object recognition under clutter**

# Effect of clutter on detection



recognition without attention

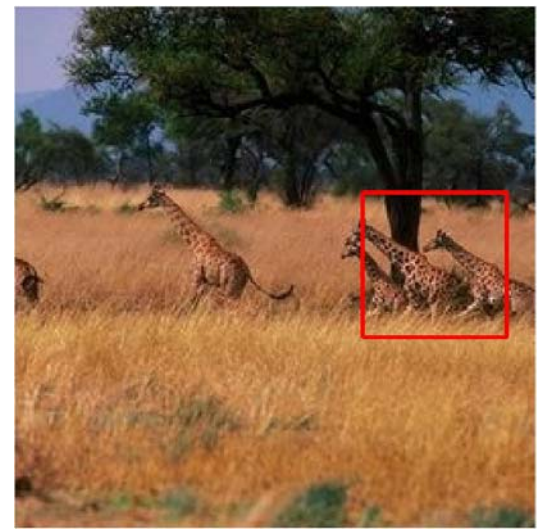
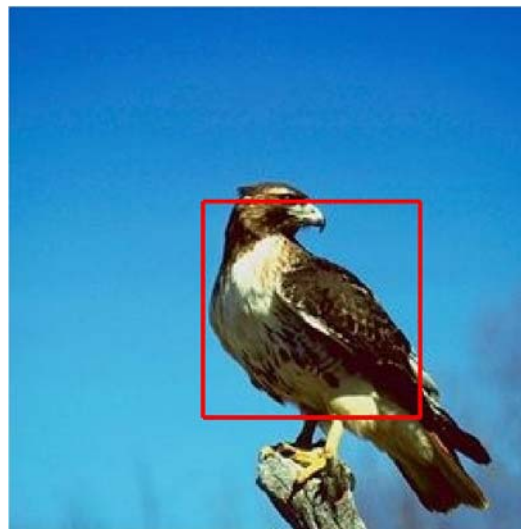
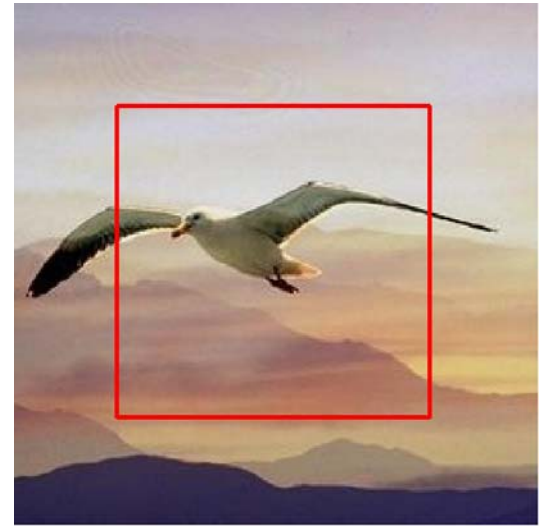
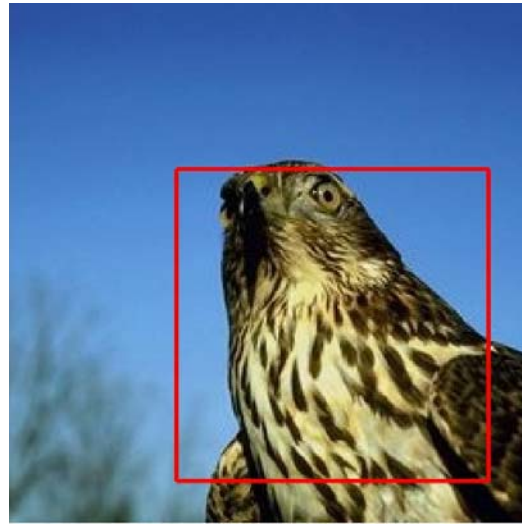


recognition under attention

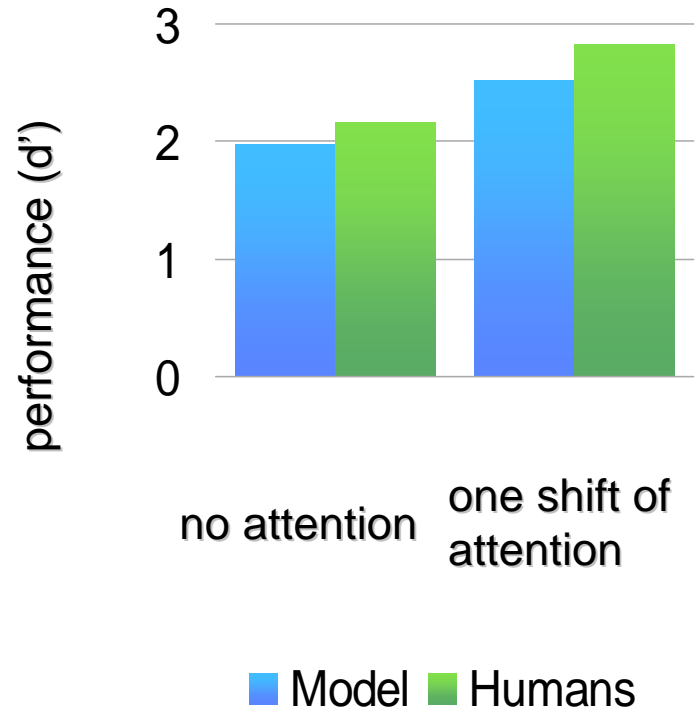
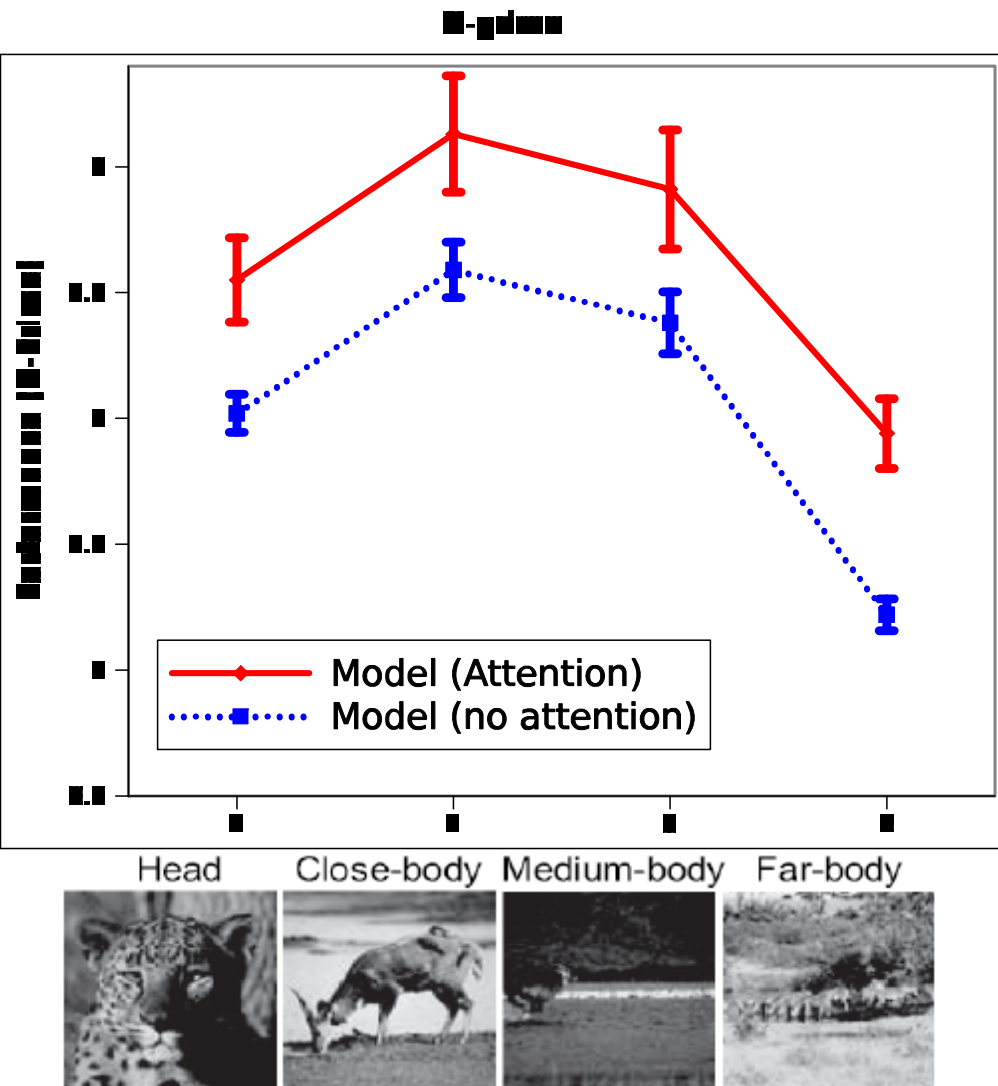
Head      Close-body      Medium-body      Far-body



# Scale and location prediction



# Performance improves under attention



Questions?

Thank you!