Putting Visual Object Recognition in Context

Mengmi Zhang[1,2], Claire Tseng[3], Gabriel Kreiman[1,2]

[1] Children's Hospital, Harvard Medical School, Boston, MA, USA

[2] Center for Brains, Minds and Machines, Harvard and MIT, Cambridge, MA, USA [3] Harvard College, Harvard University, Cambridge, MA, USA

CENTER FOR Brains Minds+ Machines

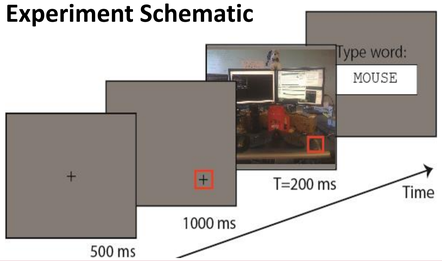
INTRODUCTION

Both humans and models rely on context for visual recognition.

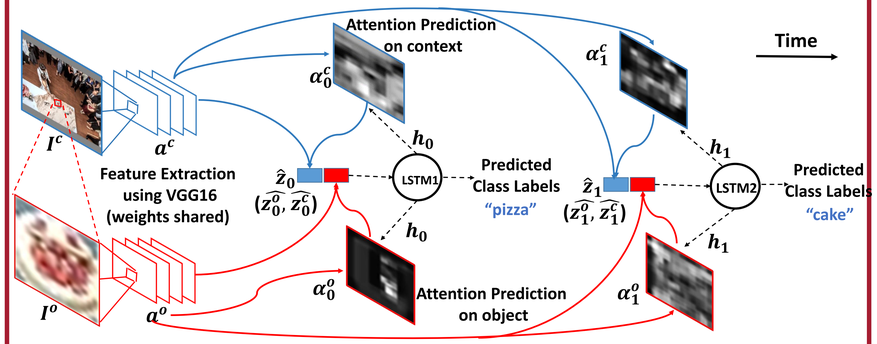
Highlights: Human benchmark + Model evaluation without re-training



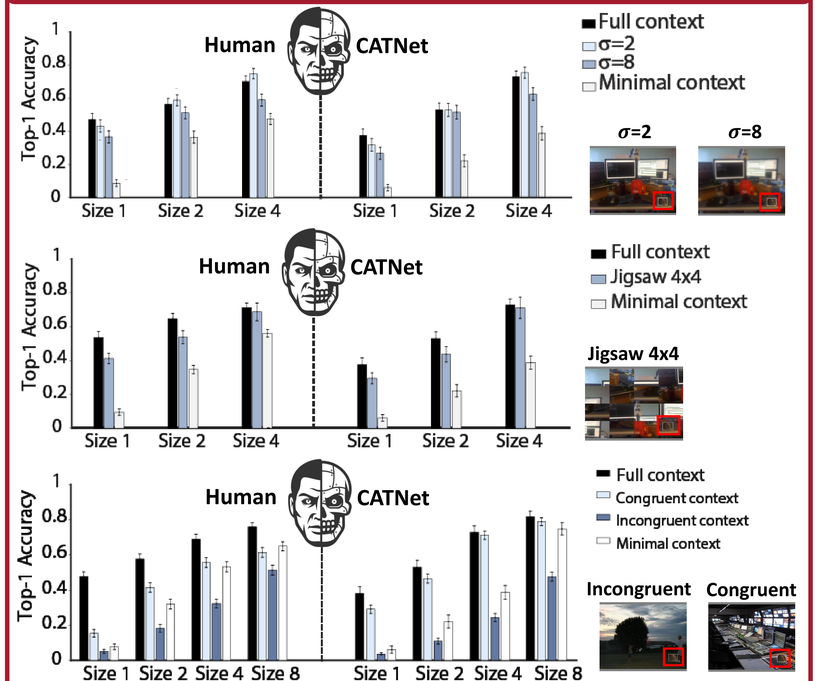
Experiment Schematic



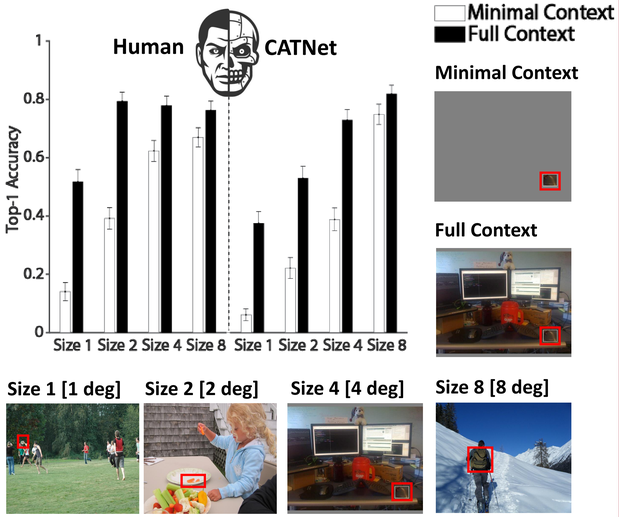
Context-Aware Two-stream Attention Net (CATNET)



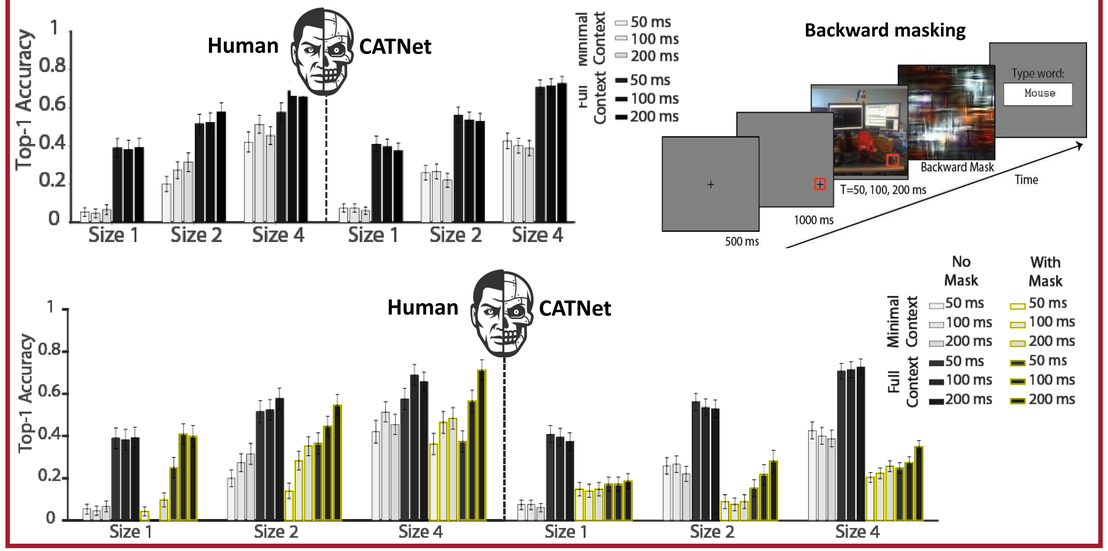
Context spatial configuration impacts recognition



The amount of context matters



Temporal dynamics of contextual modulation involves recurrence



CONCLUSIONS

Object and context size matter.

Blurred context is sufficient for recognition.

Contextual effects rely on spatial configuration.

Contextual modulation involves recurrence and happens fast.

Contact: Mengmi.Zhang@childrens.harvard.edu

CVPR Paper, source code, data

https://github.com/kreimanlab/Put-In-Context

NSF

Agency for Science, Technology and Research SINGAPORE