

# When Pigs Fly: Contextual Reasoning in Synthetic and Natural Scenes



Philipp Bomatter<sup>1,\*</sup>, Mengmi Zhang<sup>2,3,\*</sup>, Dimitar Karev<sup>4</sup>, Spandan Madan<sup>3,5</sup>,  
 Claire Tseng<sup>4</sup>, and Gabriel Kreiman<sup>2,3</sup>



<sup>1</sup>ETH Zürich   <sup>2</sup>Children's Hospital, Harvard Medical School   <sup>3</sup>Center for Brains, Minds and Machines  
<sup>4</sup>Harvard College, Harvard University   <sup>5</sup>School of Engineering and Applied Sciences, Harvard University   \*Equal contribution

## Introduction

Our synthetic **Out-of-Context Dataset (OCD)** allows for well-controlled and fine-grained study of different dimensions of context.

Contextual information is of fundamental importance to both human and machine vision.

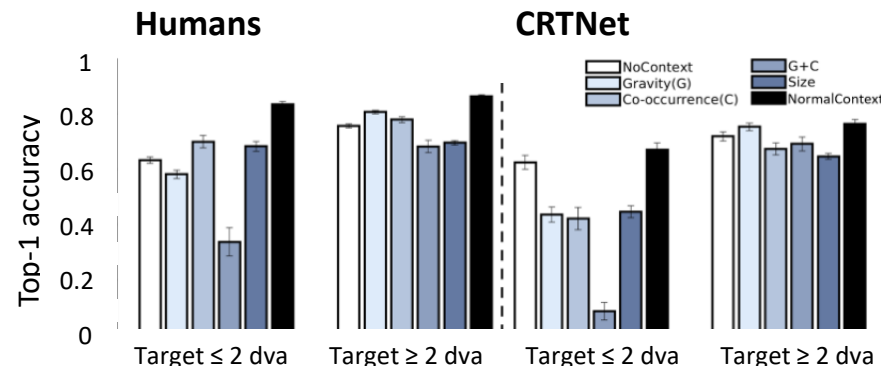
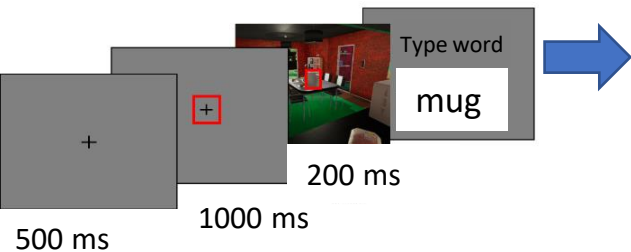
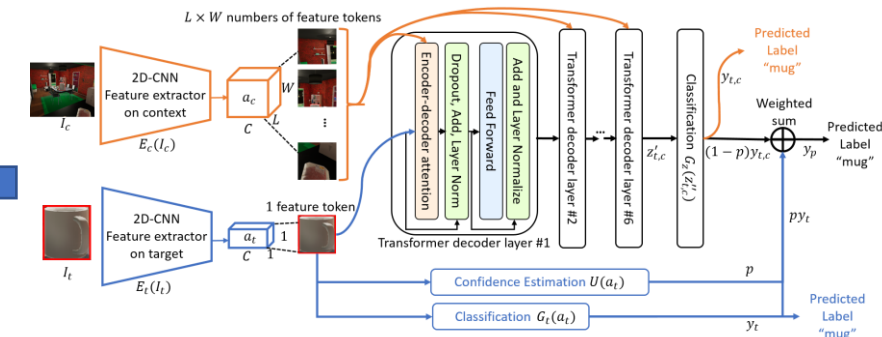
We present a new dataset with well-controlled contextual perturbations, a deep learning architecture that incorporates contextual information, and a comparison to human behavior.



**Psychophysics experiments** provide insights and an essential benchmark of human performance.

- We observe similar qualitative behavior of humans and CRTNet.
- In terms of recognition accuracy, CRTNet outperforms competitive baselines across a wide range of context conditions and datasets.

The **Context-aware Recognition Transformer (CRTNet)** leverages contextual information for object recognition.



Address correspondence to [gabriel.kreiman@tch.harvard.edu](mailto:gabriel.kreiman@tch.harvard.edu)

Source Code & Data: <https://github.com/kreimanlab/WhenPigsFlyContext>

