# Objects Are Not Scenes: Temporally Distinct Processing of Object and Scene Information

**Niklas Müller** [1,*] H. Steven Scholte [1,#] Iris I. A. Groen [2,#]

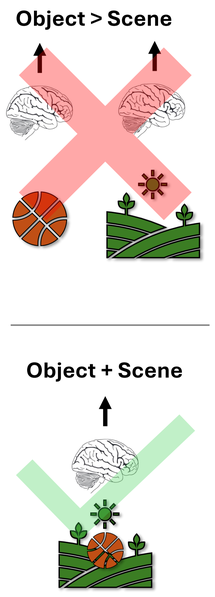
1. Psychology Research Institute 2. Informatics Institute
University of Amsterdam * **n.muller@uva.nl** # shared senior author

## Introduction

Humans process visual information fast and efficiently, enabling quick identification of their surroundings and the objects in it. It is well established that there is scene-specific processing in the human brain. A network of cortical areas processes complex visual scenes and rapidly extracts the scene gist [1,2]. It is also well established that visual cortex contains pathways that are specialized in object processing more so than scene processing [3].

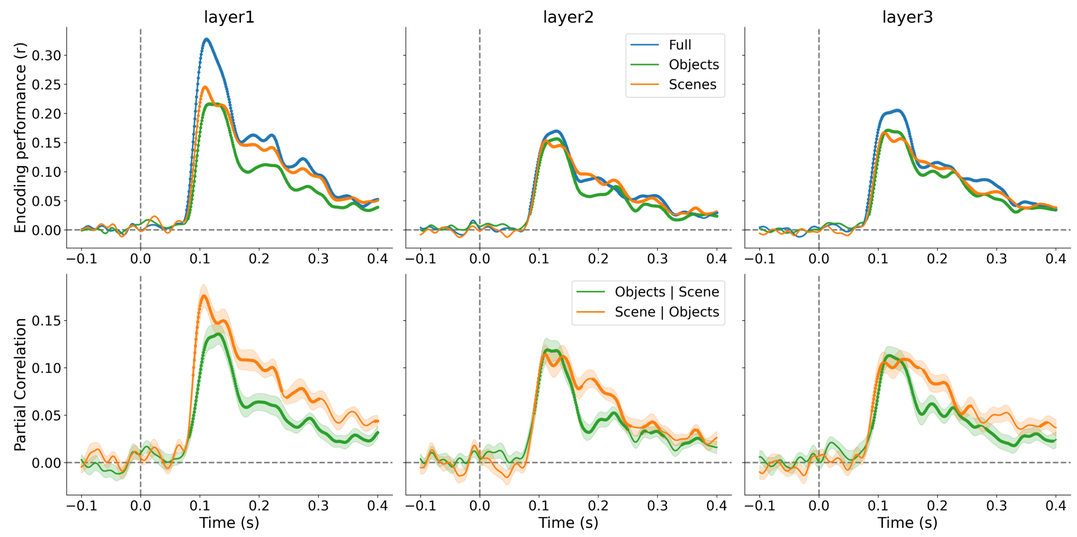
However, prior research has mostly relied on separate stimulus presentations (presenting either a scene or an object) to identify object or scene-specific processing [4,5,6]. Hence, it remains unclear whether object and scene processing are separated when viewing natural images containing both, object and scene information.

To better understand how object and scene information are processed in visual cortex, we here separate object from scene information in natural images and use Deep Neural Network (DNN) features to predict electroencephalography (EEG) recordings during natural scene viewing.




## Temporally distinct processing of object and scene information

Cross-validated encoding model performance, for **Full**, **Objects**, and **Scenes** images as input to DNN for the three maxpool layers (layer1, layer2, layer3)




**Scene information explains more unique variance and is processed earlier than object information.**

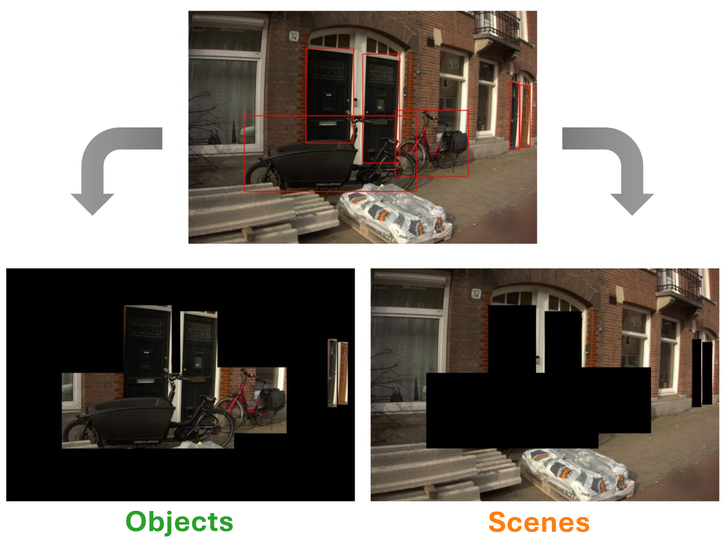
**Using later layers, object information explains more (unique) variance. However, scene information continues to explain variance even at later time points.**

## Object annotations

**Labelling Procedure**

1. Define common objects in street scenes
2. Take pictures such that selected objects occur frequently
3. Manually annotate images with bounding boxes + class labels (2/3 agreement of labellers)
4. Black out scene / object information

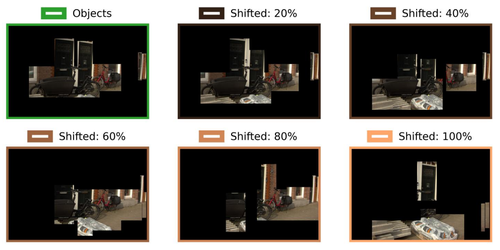
| | | |
|---|---|---|
| Compact car | Balcony door | Bollard |
| Van | Bike | Traffic sign |
| Bench | Carrier bike | Traffic light |
| Lamppost | Truck | Tree |
| Front door | SUV | |

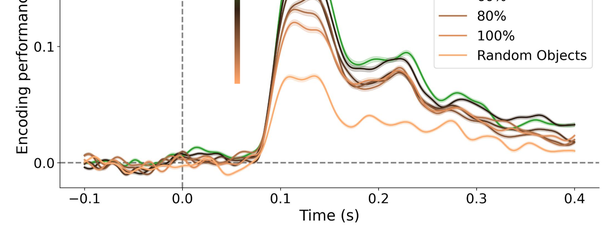




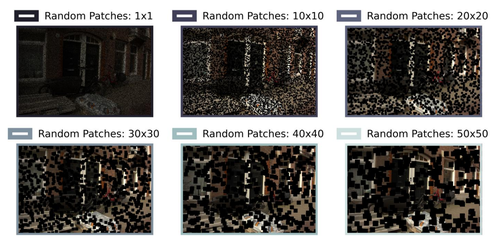

## What are objects?

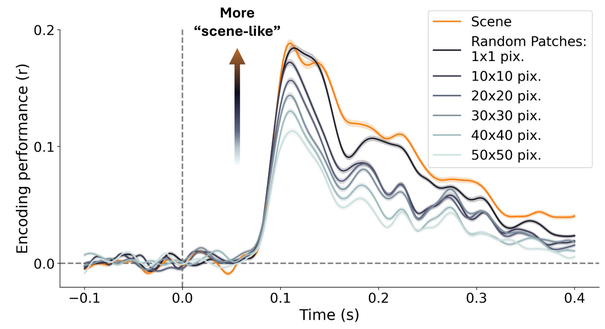
Systematically shifting object bounding boxes away from the visual object information, results in stepwise decrease of encoding performance




Using the same number of pixels while varying the scale of patches reveals rapid scene processing.




**Object and scene information, but no other spatial selection, lead to high encoding scores.**

## Encoding model

**Open Amsterdam Data Set (OADS)** [7]

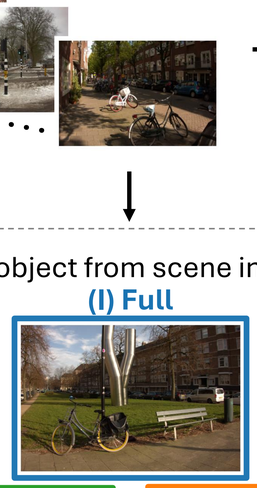
**RSVP**
**Full images**
100 ms stim
31 subjects
4680 images

Human Brain

**EEG**
(64 channels)

Separate object from scene information

**(I) Full**



**(II) Objects**

**(III) Scenes**

**Cross-Validated Encoding Model**
(per layer, subject, channel, timepoint)

AlexNet

3 Maxpool Layer Activations

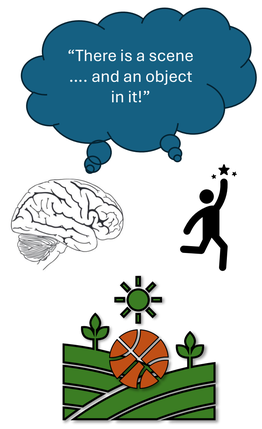
## Conclusion

These results suggest that in human visual cortex ...

1. ... objects are separately processed from scenes
2. ... scenes are processed rapidly, followed by objects
3. ... no other spatial selections are distinctly processed

This is reflected both in behaviour (collective agreement on what objects in street scenes are) as well as in neural processing (temporally distinct processing for object and scene information).

Our results show that the signatures of distinct processing pathways in visual cortex are detectable in EEG signals and inform about the temporal cascade of computations in each pathway.




**References**

[1] Oliva, *Neurobio. Att.* (2005); [2] Groen et al., *Proc. Roy. Soc.* (2017);
[3] Riesenhuber & Poggio, *Cur. Op. in Neurobio.* (2002);
[4] Epstein & Kanwisher, *Nature* (1998); [5] Grill-Spector et al., *Vis. Res.* (2001);
[6] Harel et al., *eNeuro* (2016); [7] Mueller et al, *bioRxiv (2024)*.

**Acknowledgements**
This work was funded by the Interdisciplinary PhD Programme of the Data Science Centre of the University of Amsterdam.