A single spatial transform improves predictions of neural responses by deep neural network models

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

University of Amsterdam
1. Psychology Research Institute
2. Informatics Institute

* shared senior author

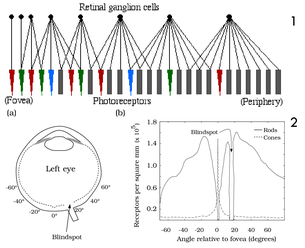
References
[1] Ohlshausen, *Sensory Processes* (2002)
[2] Wandell, *Foundations of vision* (1995)
[3] Güçlü and van Gerven, *J. NeuroSc.* (2015)
[4] Da Costa et al., *Nat. Sc. Rep.* (2024)

Acknowledgements
This work was funded by the Interdisciplinary PhD Program from the UvA Data Science Center

Modeling Human Visual Processing

- The visual field of humans is divided into foveal and peripheral regions
- Foveal and peripheral information are processed at different spatial resolutions:

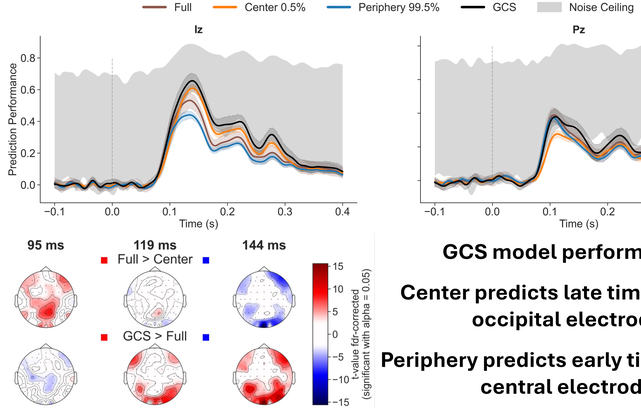
1. Foveal input is processed with high acuity and color sensitive cells
2. Peripheral input is processed with low acuity and motion sensitive cells



Encoding models using deep neural network (DNNs) features have been shown to predict neural recordings during visual processing well[3]. However, DNNs sample their visual input uniformly.

Can differential spatial sampling improve encoding model performance using DNN features?

Retinal Sampling Improves Encoding Performance



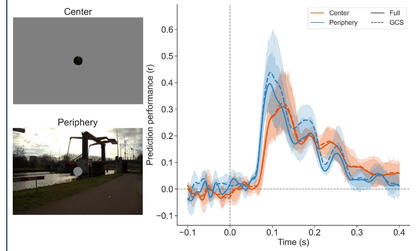
GCS model performs best

Center predicts late time points at occipital electrodes

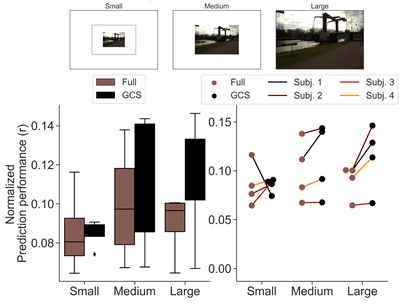
Periphery predicts early time points at central electrodes

Large field Stimulation Reveals Temporal Profiles

Additional EEG experiment (n=4) with custom stimulus conditions to test effects of selective stimulation and stimulus size.

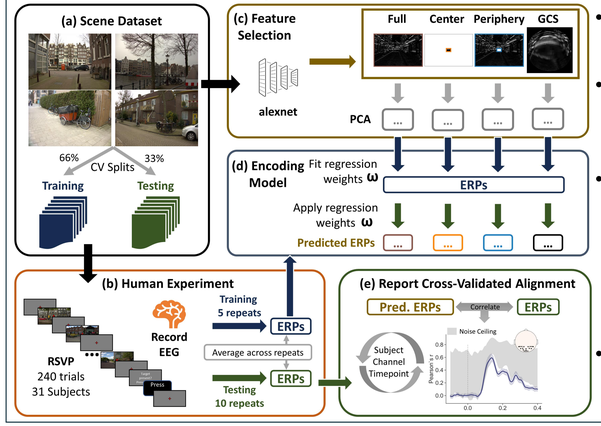


Selective visual stimulation of center versus periphery confirms temporal profiles of visual processing.



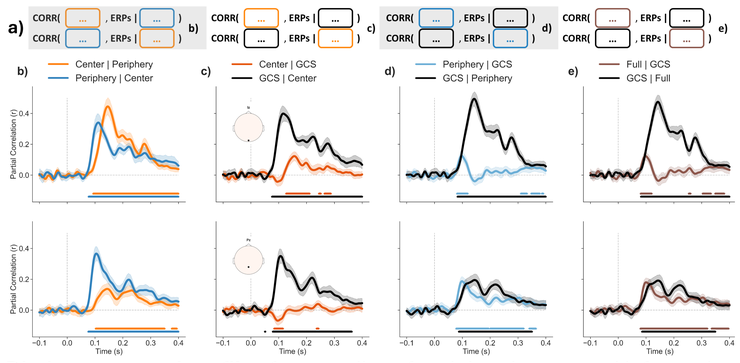
Effects of spatial sampling on EEG responses are only revealed during **large-field stimulation** that sufficiently stimulates the periphery.

Spatial Feature Selection



- High resolution stimuli
- EEG recordings from an RSVP experiment
- Spatial sampling of DNN feature maps: center and periphery crops and a ganglion cell sampling (GCS)[4]
- Cross-validated linearized encoding models

Temporal Profiles of Peripheral vs. Central Processing



Distinct temporal profiles for encoding of peripheral vs. central information
GCS model explains majority of variance at most time points and electrodes

Conclusions

- Selective spatial sampling of DNN feature maps improves encoding model performance of human EEG data
- GCS feature transform yield best performing encoding model
- Using spatial feature selection, we uncover unique temporal profiles of foveal and peripheral visual processing
- Selective spatial stimulation confirms the temporal profiles
- Importance of retinal sampling becomes apparent only when sufficiently stimulating peripheral regions

Differential spatial sampling of DNN feature maps in encoding models supports coarse-to-fine visual perception in which global, peripheral information precedes central, detailed information.