# Sample, Don't Assume: Unconstrained Receptive Field Mapping

**Niklas Müller** [1,*] Paolo Papale [3] H. Steven Scholte [1,#] Iris I. A. Groen [2,#]

A.D. de Groot Center for Cognitive AI

1. Psychology Research Institute
2. Informatics Institute
3. Netherlands Institute for Neuroscience

University of Amsterdam
* **n.muller@uva.nl**
# shared senior author

University of Amsterdam

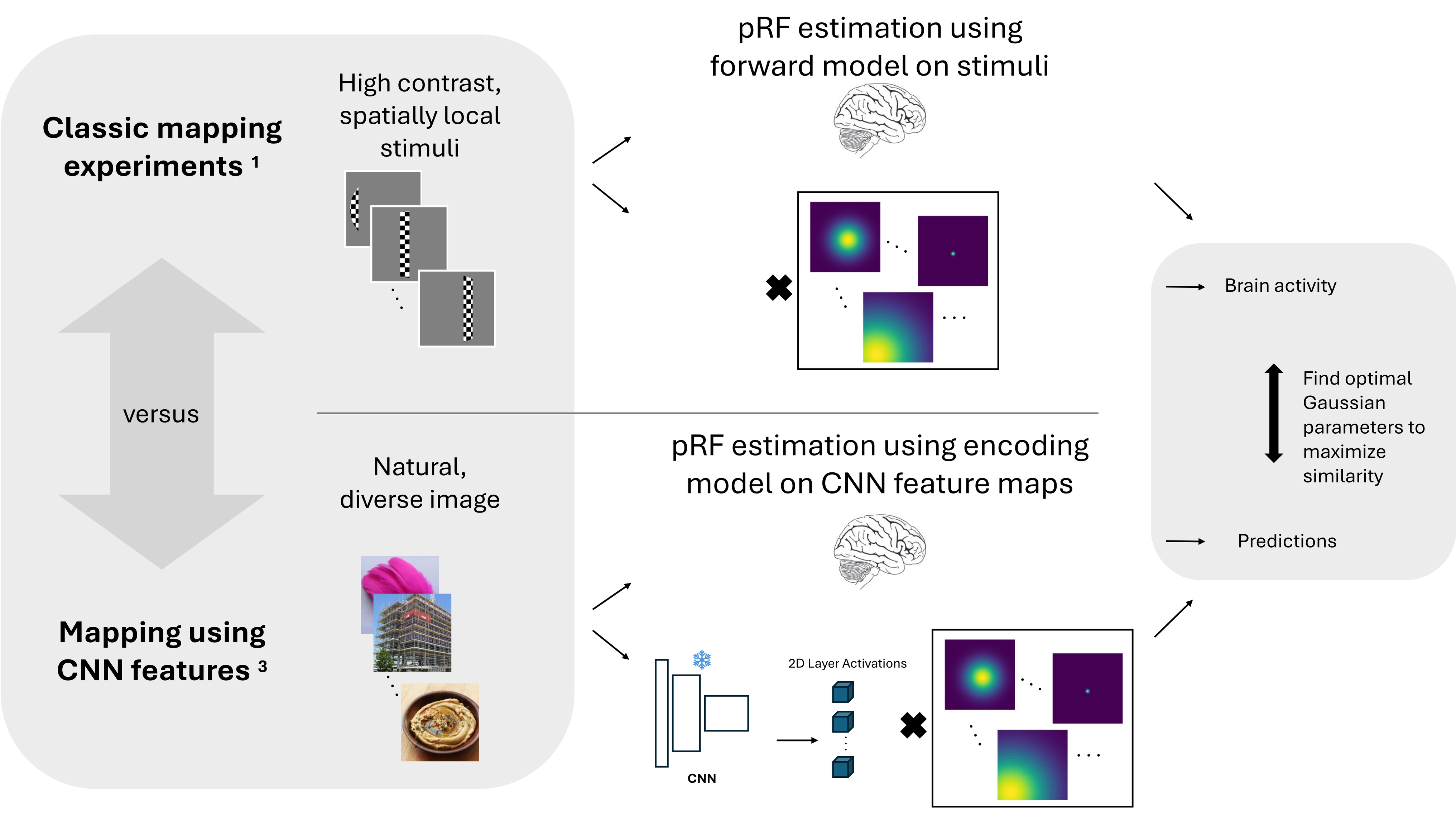
## Population Receptive Field mapping

The spatial tuning of populations of neurons, as reflected in their **population receptive field (pRF)**, is a fundamental property determining visual neural responses.

pRF geometry is typically modeled as a 2D isotropic Gaussian [1], assuming the pRF samples a circular "aperture" in the visual field. This assumption has allowed to formalize mathematical models of neural spatial tuning.

However, it has been found that using a more complex geometry can improve neural predictions [2]. **Thus, it remains unclear what assumptions to make about the geometry of pRFs.**

Advances in using Convolutional Neural Network (CNN) features [3] allow for estimating pRFs on natural images and to model instead of measure the neural spatial tuning.



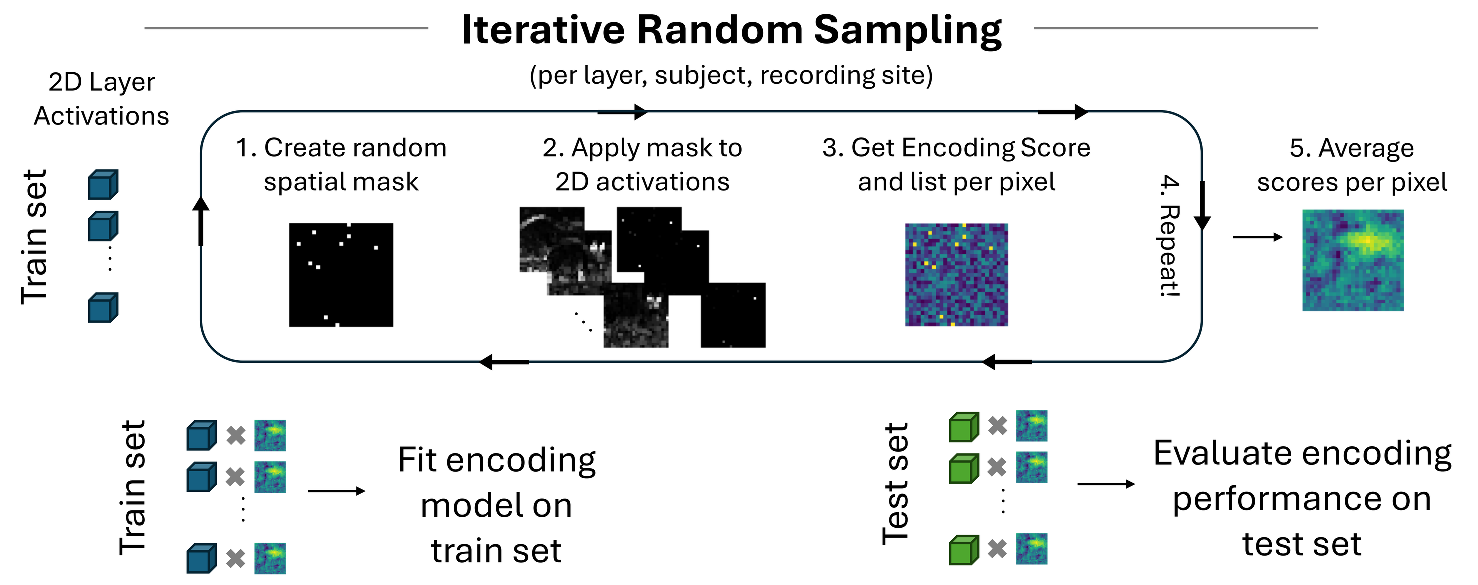

Both operate in the parameter space of a pre-defined geometry (e.g., a 2-D isotropic Gaussian) and are therefore restricted by the expressiveness of that geometry.

**What happens if we let go of the Gaussian assumption and estimate pRFs with unconstrained geometries?**

## Unconstrained Receptive Fields

Our primary goal is to estimate spatial tuning, without making assumptions about the pRF geometry. Instead of estimating pRFs in the parameter-space of e.g., a 2-D Gaussian, we **estimate pRFs in pixel-space**, by quantifying how much each pixel contributes to the encoding performance per unit.

In each iteration, we select a random subset of pixels and evaluate an encoding model that only uses visual information from the remaining pixels. The resulting encoding performance (r) is then credited to the subset of pixels and averaged per pixel across many iterations.



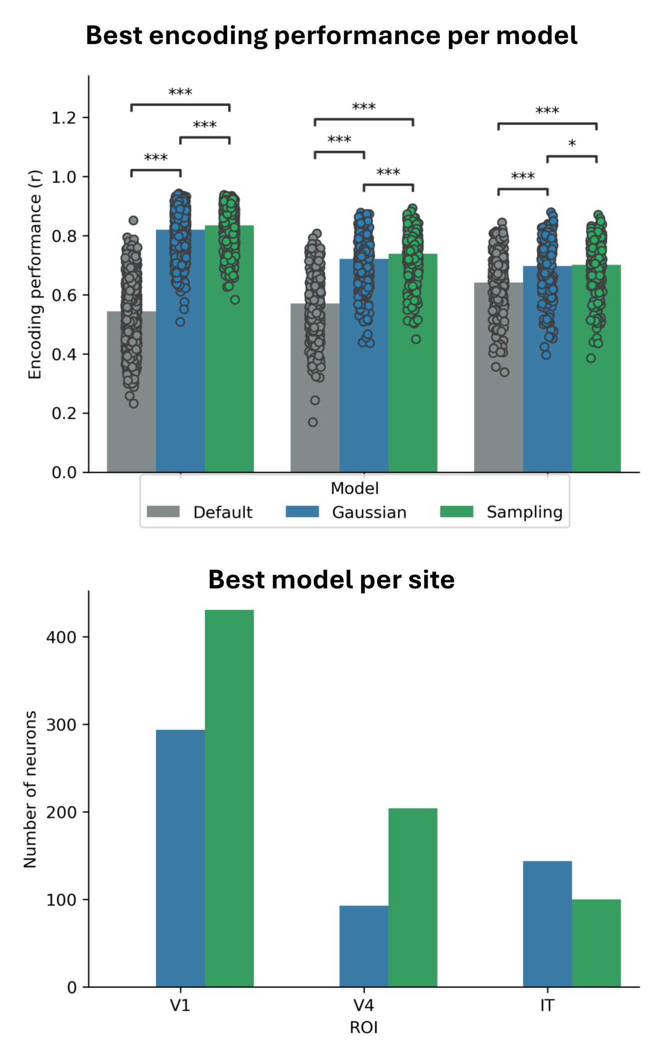

**Dataset**
We use the THINGS ventral stream spiking dataset (TVSD) [4] containing electrophysiology recordings of visual regions V1, V4, and IT of two macaque monkeys viewing natural object images from the THINGS [5] dataset.

**CNN Model**
We use the three max-pool layers of an ImageNet-trained AlexNet for feature-extraction for all encoding models.

**Spatial weighting**
We compare encoding models using different multiplicative spatial weights of feature maps: Default (equal weight at all locations), Gaussian (2-D isotropic), and the sampled pRFs. After spatial weighting, each feature map is averaged across the spatial dimensions.

**Encoding model**
For each layer, we regress the collapsed features onto the Ephys data per recording site on the train set. Finally, we evaluate encoding performance as the Pearson correlation between the predictions on the test set and the measured neural data per recording site.

## Sampling outperforms Gaussian




Using the sampled pRFs as spatial weights outperforms the best Gaussian fit across all three ROIs.

Both models consistently outperform the Default model.

While encoding performance for the Sampling and Gaussian model decreases towards higher level visual cortex, it increases for the Default model.

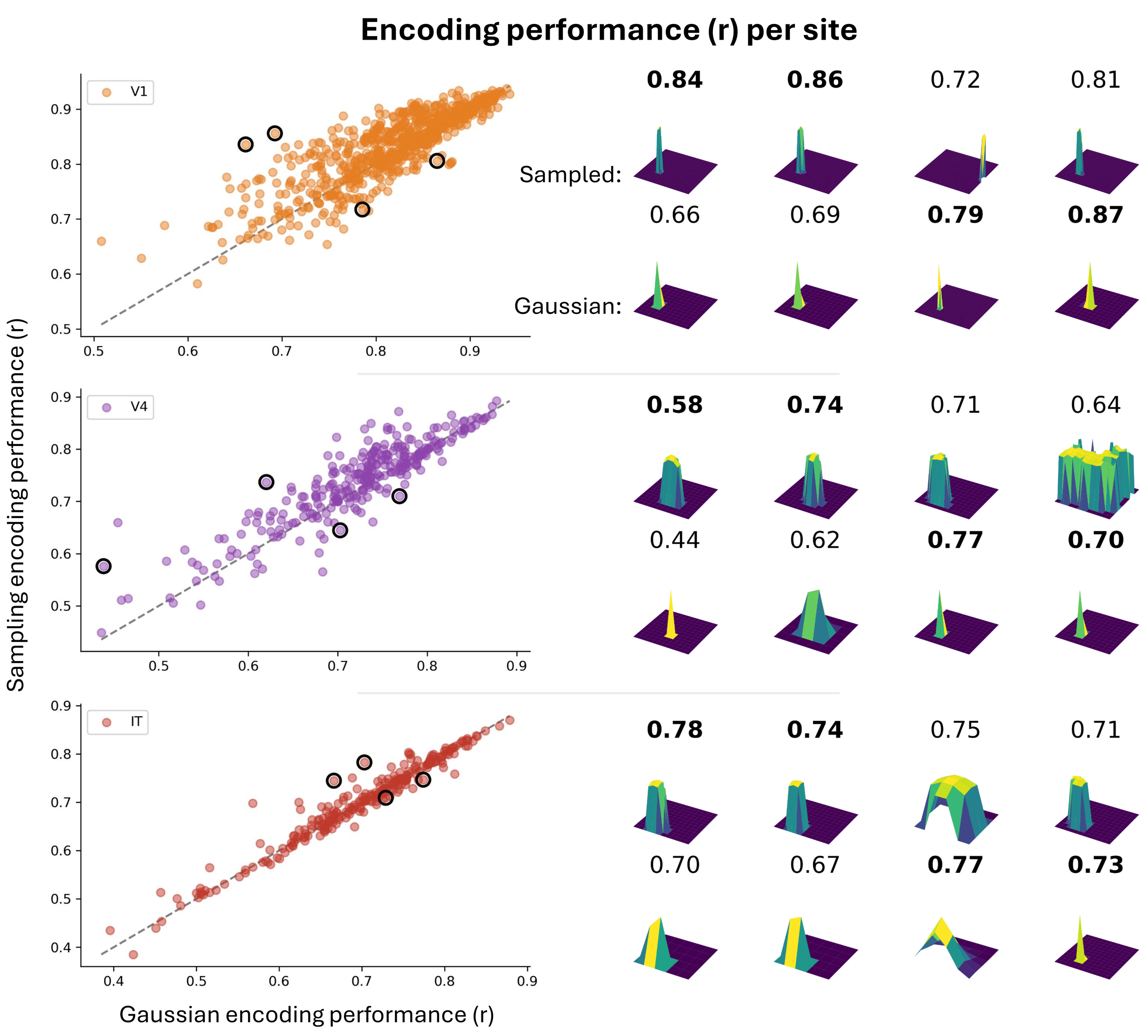
Sampling needs ~75% less computation time than Gaussian fitting.

## Receptive Field Geometry

Without making assumptions about the geometry of pRFs, we find that:

1. The sampled pRFs **resemble a wide range of tuning profile, in many cases outperforming the best Gaussian fit**
2. Very non-Gaussian looking pRFs can also achieve high encoding performance (even if the Gaussian model performs better)

Overall, the spread in encoding performance differences decreases from V1 to V4 and IT, suggesting that the Gaussian model is generally better for recording sites in higher-level visual cortex.




**Future directions**
Using this method we can explore how accurately the Gaussian (and any other geometry) can describe pRFs. More specifically, we can investigate **whether more spatially-complex tuning profiles exist** (e.g., bimodal, non-convex, non-isotropic, etc.).

Future steps would include sampling on the CNN input image instead of on the feature maps, to **increase spatial specificity and spatial resolution**. However, doing this would also interfere with the processing within the CNN itself.

## Conclusions

- Using **random sampling** of convolutional feature maps, we recovered the spatial tuning of population of neurons in monkey electrophysiology data.
- The resulting sampled pRFs are obtained in **a fully data-driven, assumption-free approach** and therefore advance our understanding of the true nature of pRFs.
- Using the sampled pRFs as spatial weights in an encoding model results in the overall **highest encoding performance**.
- Obtaining the sampled pRFs is **computationally much faster** than Gaussian fitting.

**References**
[1] Dumoulin & Wandell, *Neuroimage* (2008); [2] Aqil et al., *PNAS* (2021);
[3] St-Yves & Naselaris, *NeuroImage* (2018); [4] Papale et al., *Neuron* (2025);
[5] Hebert et al., PLOS One (2019).

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