

Region ID	Region Name	Area
V1	Primary Visual Cortex	Primary Visual
V2	Second Visual Area	Early Visual
V3	Third Visual Area	Early Visual
V4	Fourth Visual Area	Early Visual
V8	Eighth Visual Area	Ventral Stream
FFC	Fusiform Face Complex	Ventral Stream
PIT	Posterior Infero-Temporal Complex	Ventral Stream
VMV1	Ventro-Medial Visual Area 1	Ventral Stream
VMV3	Ventro-Medial Visual Area 3	Ventral Stream
VMV2	Ventro-Medial Visual Area 2	Ventral Stream
VVC	Ventral Visual Complex	Ventral Stream
V6	Sixth Visual Area	Dorsal Stream
V3A	Area V3A	Dorsal Stream
V7	Seventh Visual Area	Dorsal Stream
IPS1	Intra-Parietal Sulcus Area 1	Dorsal Stream
V3B	Area V3B	Dorsal Stream
V6A	Area V6A	Dorsal Stream
MST	Medial Superior Temporal Area	Lateral Stream
LO1	Area Lateral Occipital 1	Lateral Stream
LO2	Area Lateral Occipital 2	Lateral Stream
MT	Middle Temporal Area	Lateral Stream
PH	Parahippocampal Area	Lateral Stream
V4t	Area V4t	Lateral Stream
FST	Fundus of the Superior Temporal Sulcus	Lateral Stream
V3CD	Area V3CD	Lateral Stream
LO3	Area Lateral Occipital 3	Lateral Stream

Table A2: Visual cortex regions and areas

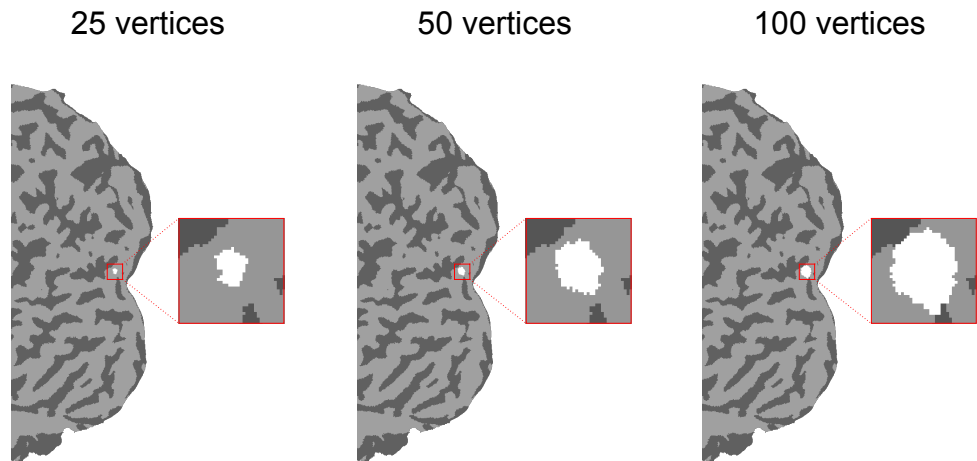


Figure A2: **Representation of the size of the searchlight on the cortical flatmap for Subject 1 from the THINGS fMRI dataset: 25, 50, and 100 vertices.**

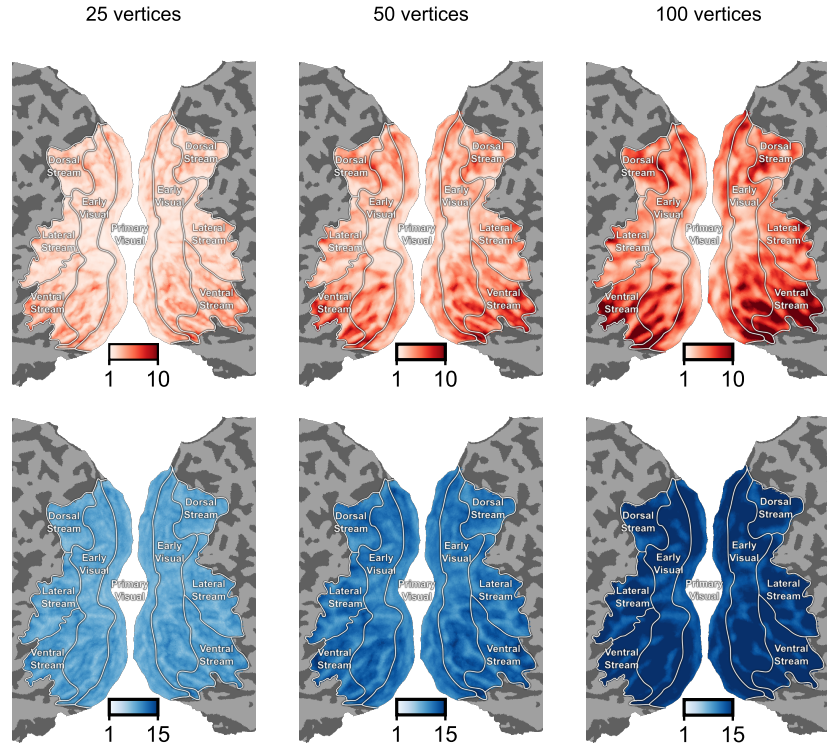


Figure A3: **Lower dimensionality with smaller neighborhood sizes.** For the neighborhood sizes shown in Fig. [A2](#), we again compute the dimensionality across the whole visual cortex, and observe similar patterns as reported in the main text.

Vertices	Left Hemisphere	Right Hemisphere
25	4.97 ± 0.05 mm	4.96 ± 0.05 mm
50	7.06 ± 0.09 mm	7.04 ± 0.08 mm
100	10.01 ± 0.12 mm	9.97 ± 0.12 mm

Table A4: Average size (in mm) and standard deviation of the searchlight area using 25, 50, and 100 vertices.

A1.5 Across subject comparison

While we report only subject 1 of the THINGS database in the main text, we here also show the results for the other two THINGS subjects, as well as all subjects in the BOLD5000 (Figs. A4, A5). Across all, we observe very similar trends in increases in dimensionality along the visual processing hierarchy of the ventral visual cortex.

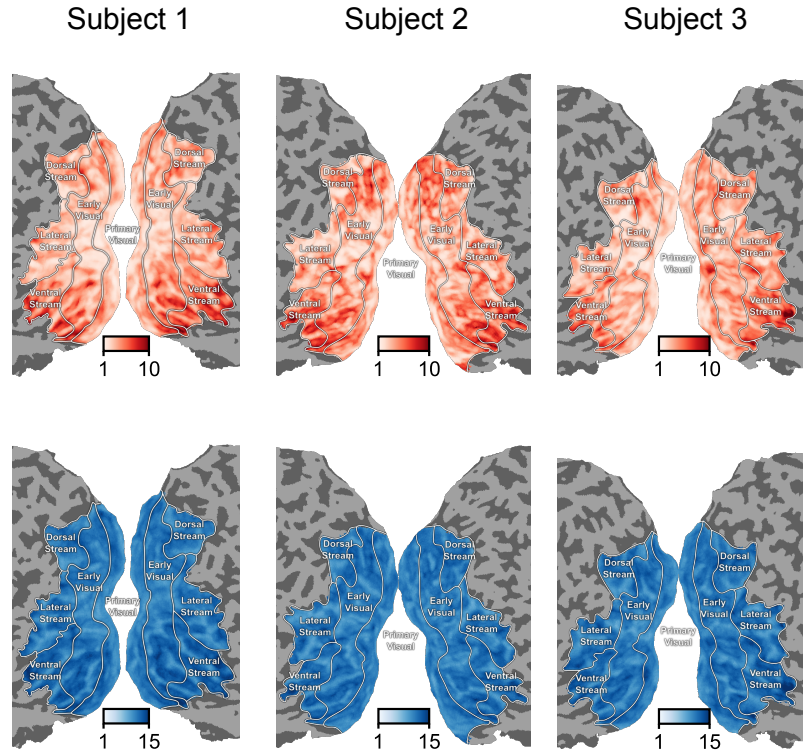


Figure A4: Effective and intrinsic dimensionality for all three subjects of the THINGS database using 50 vertices.

A1.6 Dimensionality in the left and right hemisphere

Since the two hemispheres contain different numbers of vertices, we compute the correlation of sampled dimensionality across regions (see Fig. 2b) by averaging over 100 randomly selected vertices per region, repeated 100 times. We observe that the hemispheres in the visual cortex are quite similar, they process slightly different parts of the visual field. We therefore here confirmed that across all subjects and datasets, the dimensionality measures are similar within the left and the right hemisphere (Figs. A6, A7).

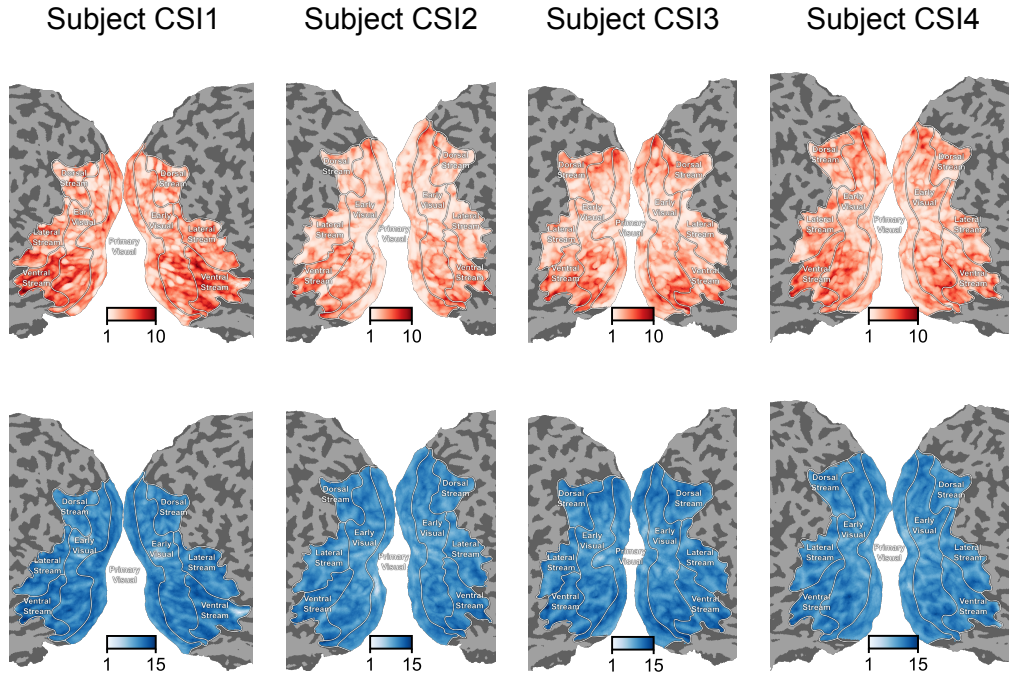


Figure A5: Effective and intrinsic dimensionality for all three subjects of the BOLD5000 database using 50 vertices.

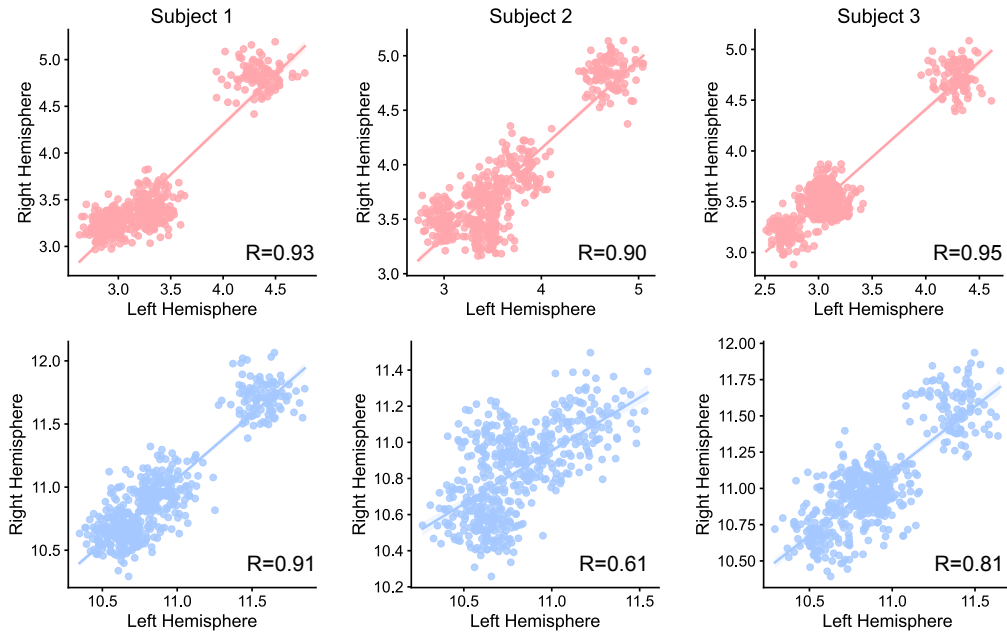


Figure A6: Correlation between the dimensionality of the left and right hemisphere across brain regions for all THINGS subjects. Top: effective dimensionality, bottom: intrinsic dimensionality.

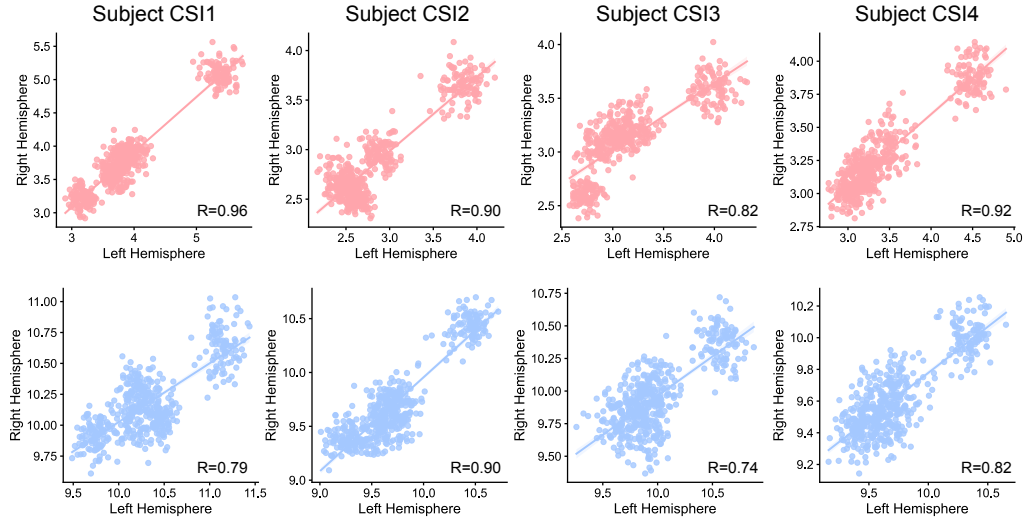


Figure A7: Correlation between the dimensionality of the left and right hemisphere across brain regions for all BOLD5000 subjects. Top: effective dimensionality, bottom: intrinsic dimensionality.

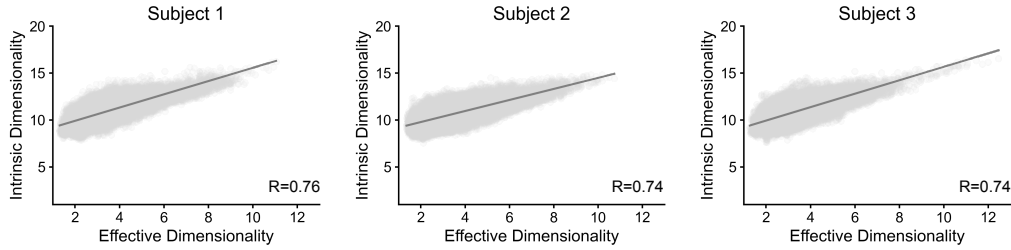


Figure A8: Correlation between effective and intrinsic dimensionality of vertices for THINGS subjects.

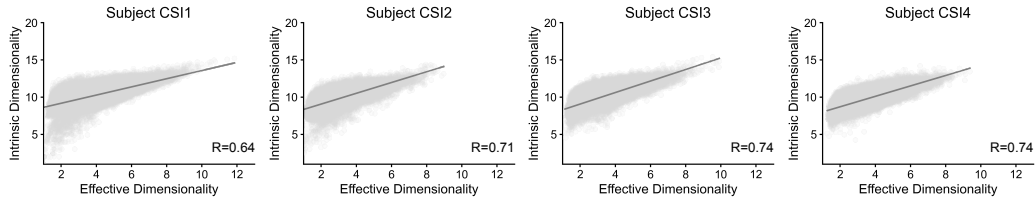


Figure A9: Correlation between effective and intrinsic dimensionality of vertices for BOLD5000 subjects.

483 A1.7 Effective vs. Intrinsic dimensionality in the brain

484 We observe highly correlated values for the linear (effective) and non-linear (intrinsic) dimensionality
 485 estimation in the visual cortex for both data sets and subjects (Figs. A8, A9). This is different to
 486 the ANNs, where for the layers with highest effective dimensionality (later layers), the intrinsic
 487 dimensionality dropped (cf. Fig. 2).

A1.8 Random sampling in fMRI

In the main text, we report results using a searchlight approach to compute the dimensionality. To test whether the differences in dimensionality are effected or even driven by the correlative structure of the fMRI signal in nearby regions, we repeated the dimensionality computation sampling randomly vertex activity within an area (primary, early, ventral, lateral, dorsal). We find that even with this random sampling approach, the ventral areas have a higher dimensionality than all other areas (Fig. A10), confirming our results with the searchlight approach and suggesting that the dimensionality is only slightly affected by the spatial correlative structure of the fMRI signal.

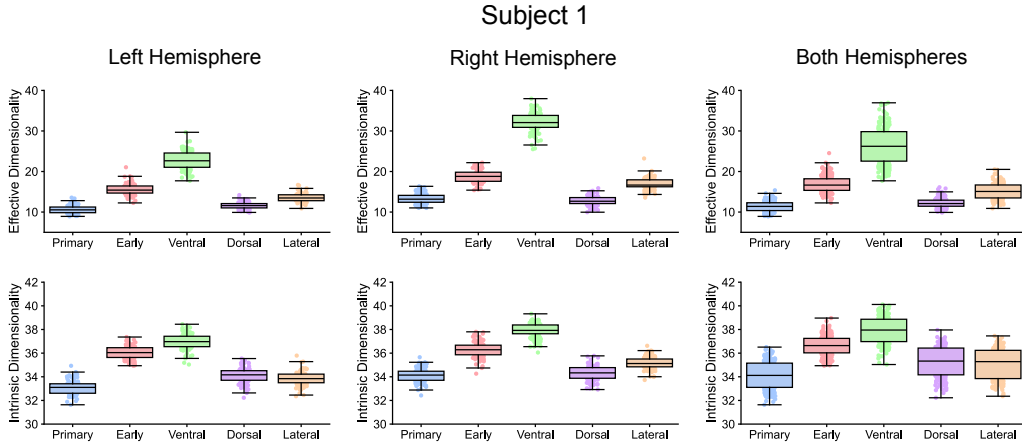


Figure A10: Calculation of dimensionality sampling random vertices of the different areas instead of using neighbors of vertices. Top: effective dimensionality, bottom: intrinsic dimensionality.

A1.9 Superlinear dimensionality increase in the ventral stream

We further quantified the increase in dimensionality along the processing hierarchy in the ventral stream from primary, to early, to the ventral areas. We quantified the hierarchy using the average geodesic distance from the center of the primary visual cortex. Generally, we find a superlinear increase along the hierarchy for all THINGS subjects (Fig. A11) and BOLD5000 subjects (Fig. A12).

A1.10 Random sampling and random projections in ANNs

How features in an ANN are preprocessed when studying brain alignment can have a significant impact on the results [30]. In dimensionality analyses, it is important to restrict the number of features across layers or regions to ensure valid comparisons. A common approach is to use random projections, which preserve the pairwise distances in the data while reducing dimensionality. While this method retains overall geometric properties, it also mixes unit activations through arbitrary linear combinations, potentially distorting unit-level selectivity, which may be critical for interpreting brain-model correspondence [30].

Previous work has linked individual ANN units to single neurons in the brain [42], supporting unit-level comparisons. However, the fMRI signal reflects a population-level average across thousands of neurons within a voxel, suggesting a mismatch in spatial scale. To address this, we adopt global average pooling, which compresses feature maps into coarse-grained, image-level representations that potentially better approximate the aggregate nature of fMRI signals. This approach preserves semantic content while discarding spatial specificity, similarly to the spatial integration performed by millimeter-scale cortical sampling in fMRI. To evaluate how preprocessing affects dimensionality estimates, we computed both effective and intrinsic dimensionality using three approaches: (i) global average pooling with feature map subsampling, (ii) randomly sampled units that preserve selectivities (unit-level features), and (iii) random projections that preserve geometric structure while reducing feature dimensionality. For doing the comparison between methods, we subsample images from

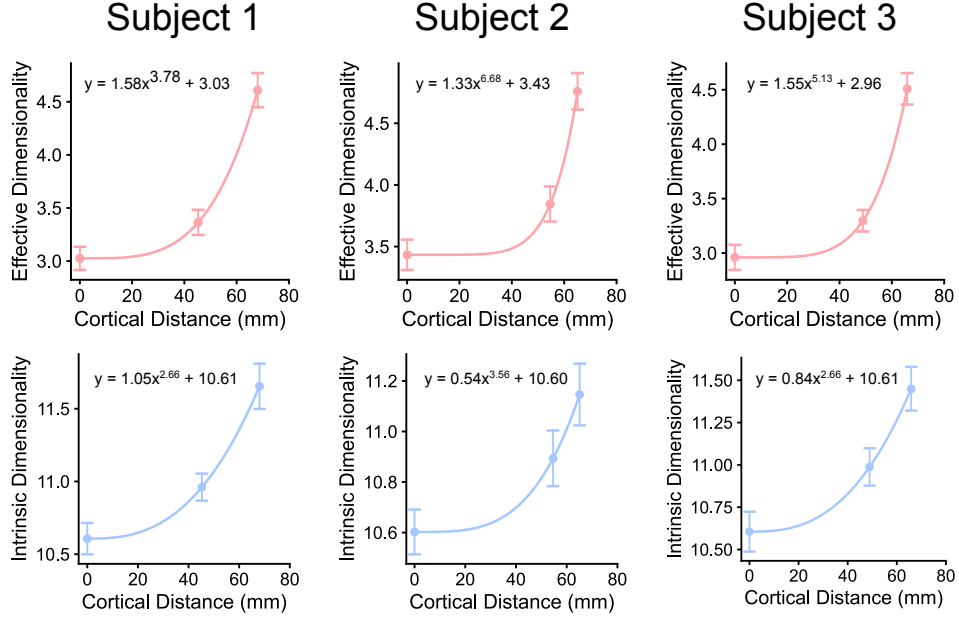


Figure A11: **Superlinear increase in effective (top) and intrinsic (bottom) dimensionality in all subjects of the THINGS database.** Cortical distance is computed as the average geodesic distance between the center of V1 to the other areas. Each dot represents the primary, early, and ventral areas, respectively.

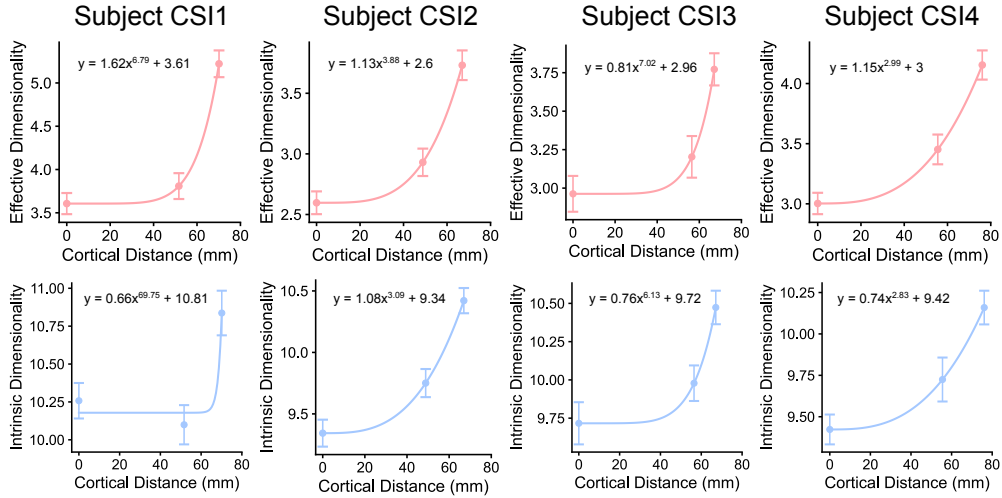


Figure A12: **Superlinear increase in effective (top) and intrinsic (bottom) dimensionality in all subjects of the BOLD5000 database.** Cortical distance is computed as the average geodesic distance between the center of V1 to the other areas. Each dot represents the primary, early, and ventral areas, respectively.

520 the dataset to reduce computational costs. We sampled 500 images from the THINGS dataset,
521 selecting only those observed by the subjects, and repeated each analysis 100 times per layer. For
522 subsampling, we selected 50 pooled feature maps or units, while for random projections, we reduced
523 the dimensionality to 50. Note that due to the smaller number of images we used, the curves obtained
524 here for global average pooling differ from those in Figure 2.

For effective dimensionality, we find that only global average pooling produces the superlinear increase across layers that mirrors the pattern observed in the ventral visual stream of the brain (Fig. A13). In contrast, unit sampling and random projections both yield the previously reported expansion-compression trend, even for effective dimensionality. For intrinsic dimensionality, this expansion-compression pattern persists across all preprocessing methods (Fig. A14), suggesting that nonlinear representational geometry in ANNs is shaped more by task constraints and training objectives than by feature selection alone.

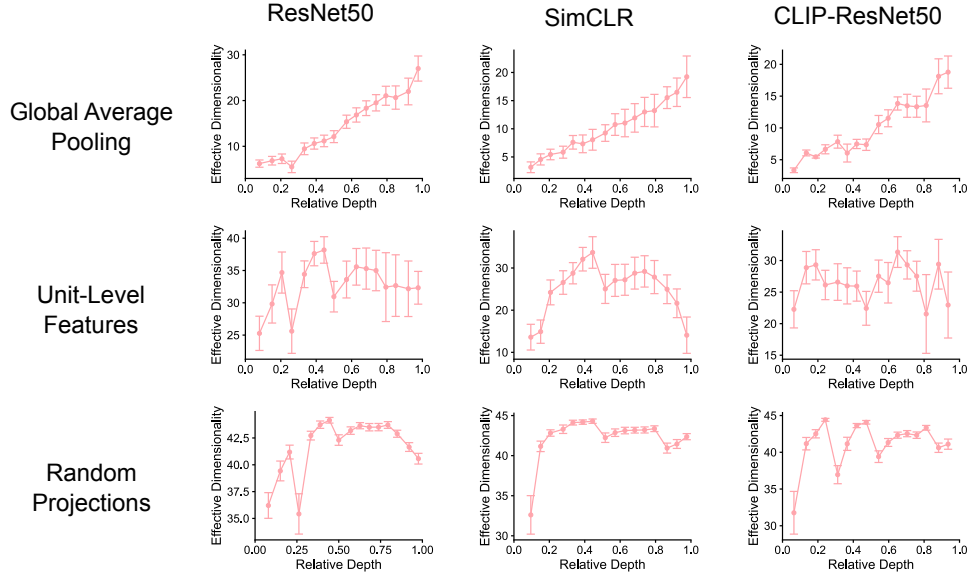


Figure A13: Effective dimensionality in ANNs using different strategies to estimate dimensionality.

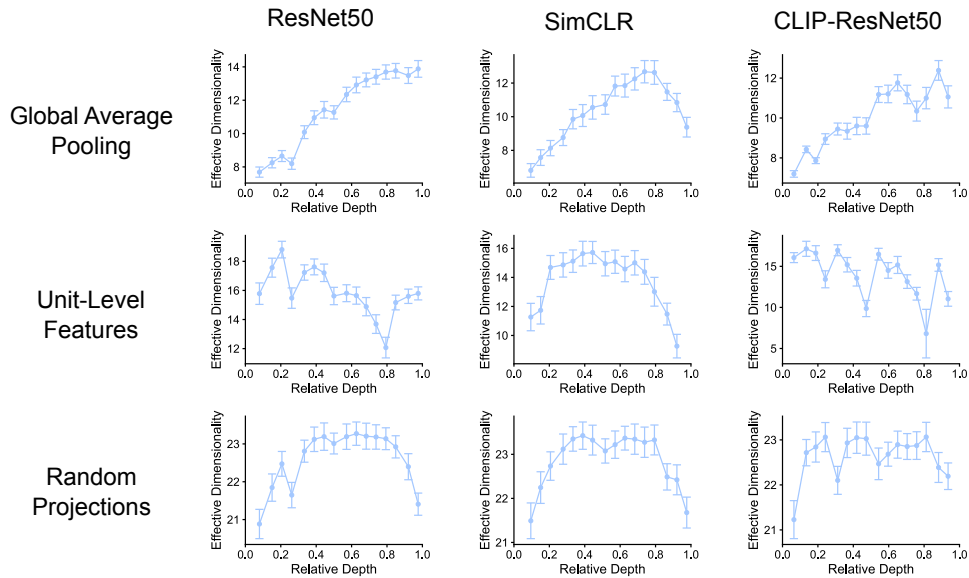


Figure A14: Intrinsic dimensionality in ANNs using different strategies to estimate dimensionality.

532 A1.11 Extracting low-level feature representation in images

533 We performed PCA on the THINGS images viewed by subjects during the experiments to extract
 534 low-level visual features [25]. Ordering the images by their loading on the first PCs reveals contrast,
 535 spatial frequency and color components (Fig. A15). The first four components already explain
 536 43.67% of the total variance, highlighting the low-dimensionality of the input data. Due to copyright
 537 restrictions, we do not display the original images. Instead, we project and show images from the
 538 THINGSplus dataset [39].

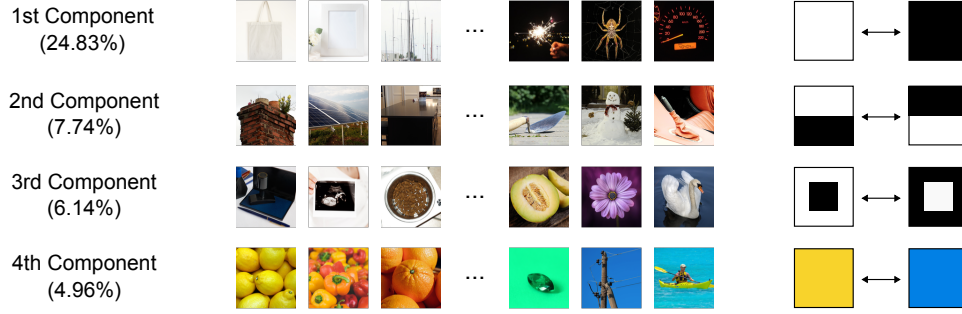


Figure A15: Image-level PCA reveals low-level features in the images.

539 A1.12 Replication of the representation of low- and high-level features in the brain

540 We repeated all prediction analyses, including stimuli, low-level image features, categories, and
 541 abstract concepts, for the other two subjects in the THINGS database. For category decoding, we
 542 selected 100 of the 720 categories observed by each subject. Each category was represented by 12
 543 different images, resulting in a total of 1,200 images. For stimulus decoding, we used a subset of 100
 544 images, each shown 12 times across sessions, also totaling 1,200 image presentations. For decoding
 545 categories we use logistic regression, and for decoding stimuli, components and concepts we use
 546 ridge regression. All decoding analyses were performed using leave-one-session-out cross-validation.
 547 For visualization purposes and to illustrate overall trends, we fit the data using the curve $y = ax^b + c$,
 548 where x represents cortical distance and y denotes either accuracy or R^2 . If the curve fitting procedure
 549 fails to converge, we instead fit a linear model $y = ax + b$.

550 We again find a reduced representation of stimuli and low-level features in later brain areas of the
 551 ventral stream (Figs. A16a, A17a, cf. Fig. 3a in the main text). The decoding performance of the
 552 categories and abstract concepts increases often superlinearly with increasing distance to the primary
 553 visual cortex (Figs. A16b, A17b, cf. Fig. 4a in the main text).

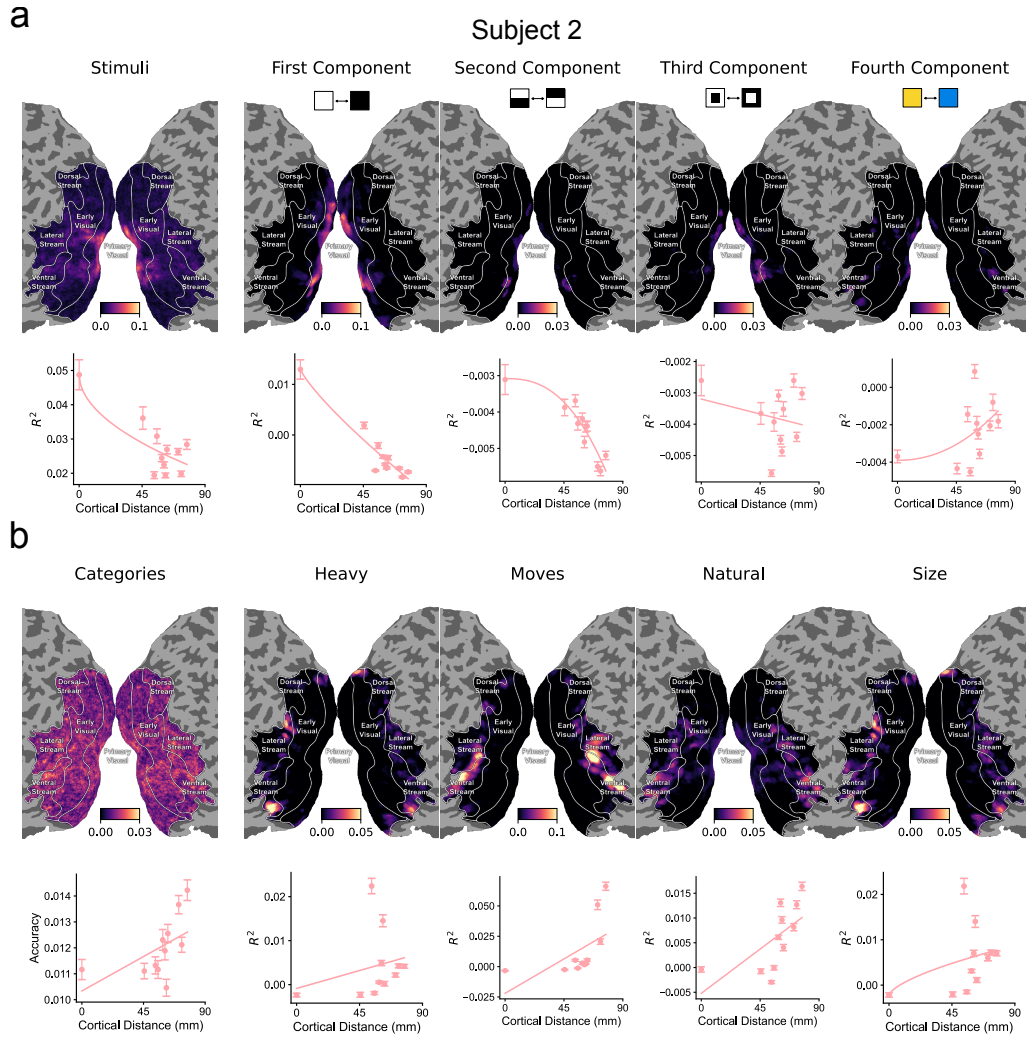


Figure A16: Low-level (a) and high-level (b) feature decoding in THINGS Subject 2.

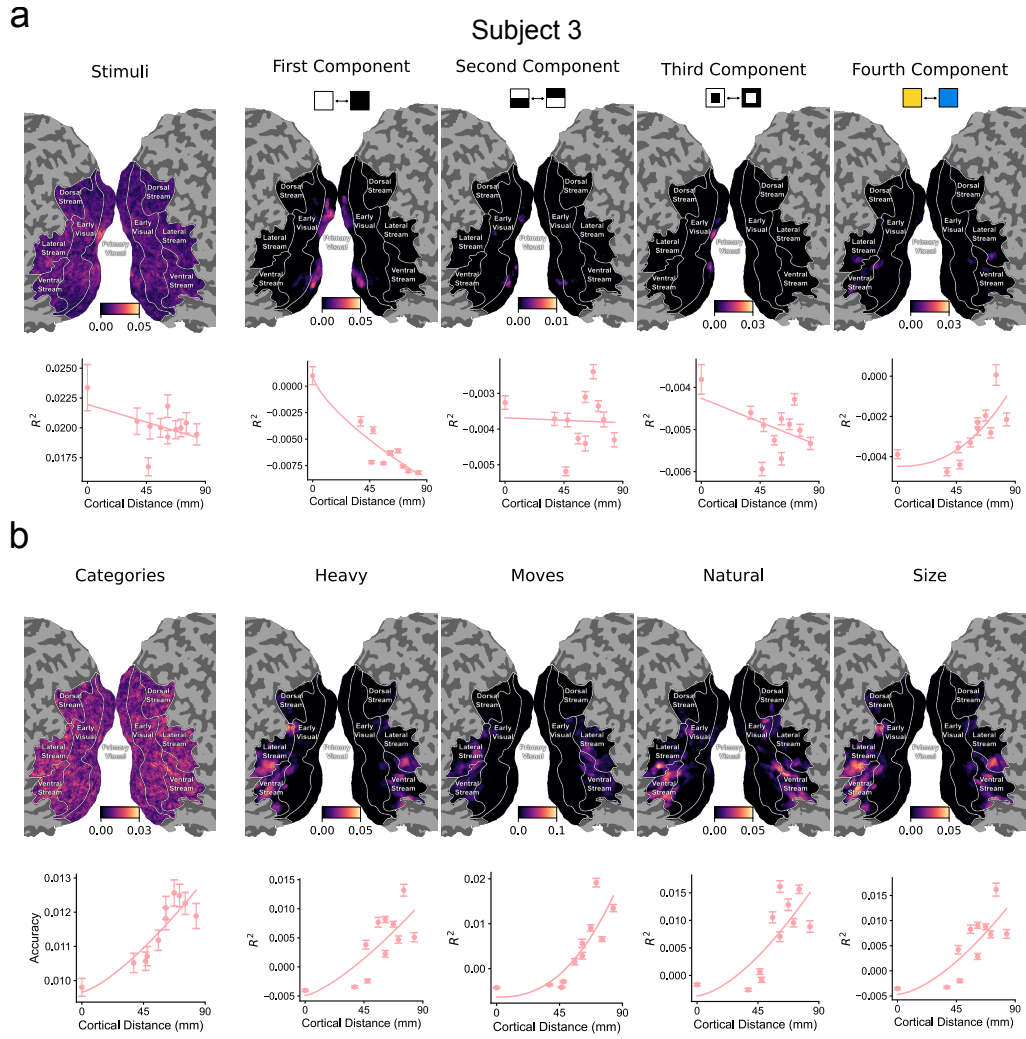


Figure A17: Low-level (a) and high-level (b) feature decoding in THINGS Subject 3.