Predicting brain function from anatomy using geometric deep learning

Fernanda L. Ribeiro[a,b], Steffen Bollmann[c], Alexander M. Puckett[a,b]

[a]School of Psychology, The University of Queensland; Brisbane QLD 4072; Australia

[b]Queensland Brain Institute, The University of Queensland; Brisbane QLD 4072; Australia

[c]Centre for Advanced Imaging, The University of Queensland; Brisbane QLD 4072; Australia

Corresponding author:

Fernanda L. Ribeiro

School of Psychology

University of Queensland

St. Lucia QLD 4072

Australia

Email: fernanda.ribeiro@uq.edu.au

Phone: +61432643672

Abstract

Whether it be in a single neuron or a more complex biological system like the human brain, form and function are often directly related. The functional organization of human visual cortex, for instance, is tightly coupled with the underlying anatomy. This is seen in properties such as cortical magnification (i.e., there is more cortex dedicated to processing foveal vs. peripheral information) as well as in the presence, placement, and connectivity of multiple visual areas – which is critical for the hierarchical processing underpinning the rich experience of human vision. Here we developed a geometric deep learning model capable of exploiting the actual structure of the cortex to learn the complex relationship between brain function and anatomy in human visual cortex. We show that our neural network was not only able to predict the functional organization throughout the visual cortical hierarchy, but that it was also able to predict nuanced variations across individuals. Although we demonstrate its utility for modeling the relationship between structure and function in human visual cortex, geometric deep learning is flexible and well-suited for a range of other applications involving data structured in non-Euclidean spaces.

Keywords

human connectome project, fMRI, 7T, vision, cortical surface, manifold, neural network, machine learning, retinotopy, visual hierarchy

1. Introduction

The modern-day study of visual perception has led to an unprecedented understanding of the neurobiological mechanisms underlying sensory information processing in the brain. Although first being of interest to those pursuing basic science, this understanding has since translated to more practical applications, notably by inspiring the development of convolutional neural networks (CNNs)[1] which have revolutionized the field of artificial intelligence[2]. In the human visual system, light reaches the retina where it is converted into electrical impulses that are transferred to the lateral geniculate nucleus and then through a hierarchy of cortical areas (Figure 1)[3,4]. At the bottom level of the hierarchy (primary visual cortex), neurons with small receptive fields (i.e., the coverage area in the visual field to which a neuron is responsive) capture simple visual features such as spatial orientation[3]. As those signals travel through higher order visual areas (comprised of neurons with increasing receptive field sizes) those simple visual features are combined, enabling complex features to be mapped from the visual input. These more complex features range from contours and textures (low to mid-level) to complete faces near the top of the hierarchy[5]. This type of hierarchical processing is imitated by CNNs by using filters with learnable parameters which are convolved across data representations with the aim of capturing common features in the input. At early layers of CNNs, those filters capture simple features, while in deeper layers more complex features are detected due to the complex combination of the inputs of previous layers. Here, we report that the story has – in a sense – come full circle. By leveraging recent advances in the application of CNNs to non-Euclidean spaces[6], we show that deep learning can be used to predict the detailed functional organization of the very brain regions that inspired them (i.e., the areas of the visual cortical hierarchy). Importantly, these

predictions are made from anatomical properties alone and vary according to each individual's specific brain anatomy.

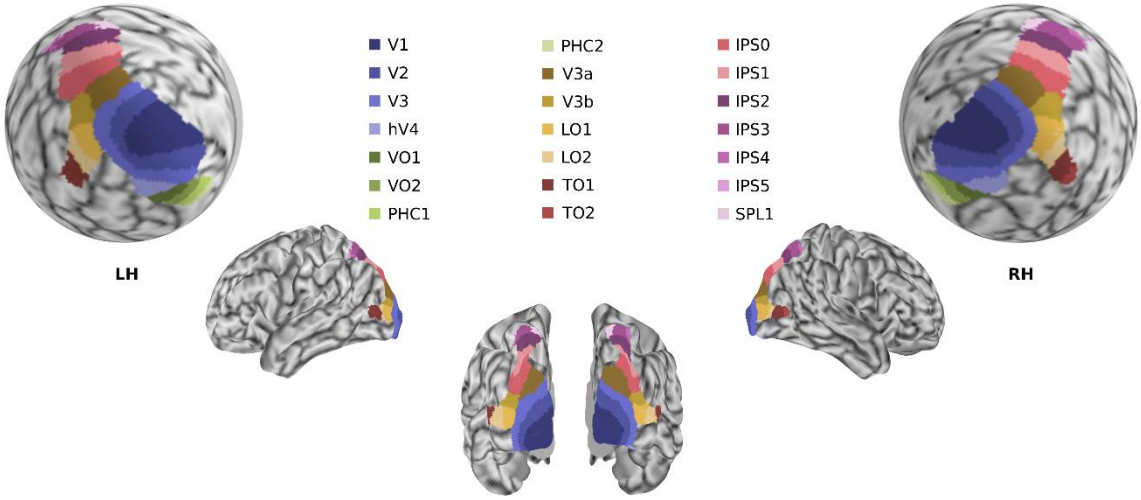


Figure 1 - Hierarchical organization of the visual cortex. The 21 atlas-based[7] visual areas in both left (LH) and right (RH) hemispheres displayed on spherical surface and fiducial surface (lateral and posterior views) models. Different colors represent different visual areas. Central legend indicates the name of each visual area.

Convolutional neural networks are conventionally applied to data represented in two- or three-dimensional Euclidean space (Figure 2a). A great deal of data, however, is not best represented in such a simple and regular domain[6]. In the field of neuroscience, for instance, data are often represented on surface models of the brain as many cortical properties only make sense considering their specific location with respect to the various sulci and gyri[8]. These surface models are essentially sheet-like meshes that are intricately folded in three-dimensional space, maintaining the geometrical structure of the brain[9]. Importantly, two points on the surface model might be close to one another in Euclidean space but far apart in terms of cortical distance following the surface (e.g., on opposing banks in a sulcus). To account for this and other non-Euclidean properties, techniques have recently been developed that generalize deep neural network models to non-Euclidean spaces such as surfaces (Figure 2b) and

graphs – with these techniques collectively being referred to as geometric deep learning[6].

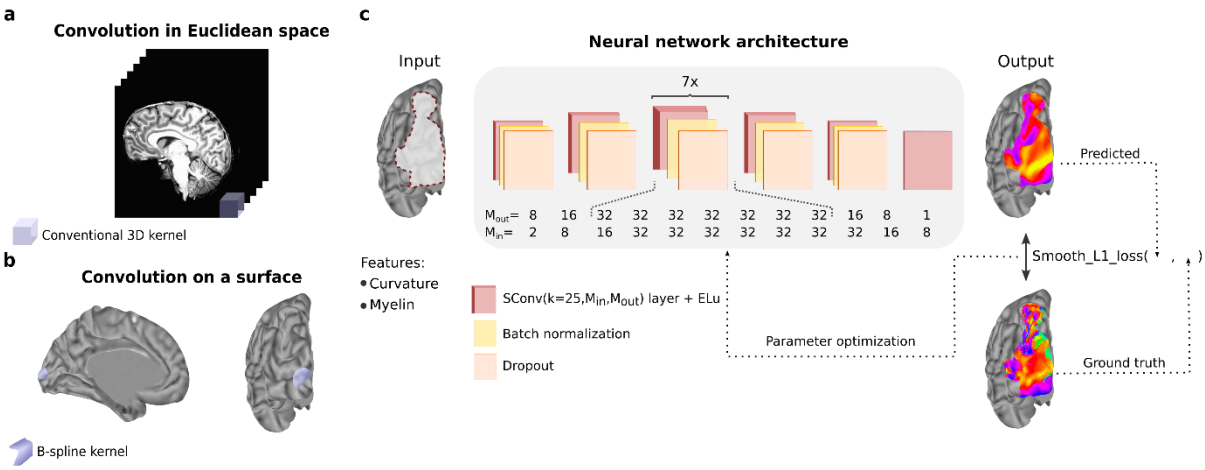


Figure 2 - Predicting brain function with geometric deep learning. a, Convolution in 3D Euclidean space involves convolving a conventional 3D kernel across multiple slices of 2D images. **b,** Convolution on a surface, a non-Euclidean space, involves convolving a B-spline kernel directly on the surface, in which the kernel is able to aggregate information in local neighborhoods. **c**, Geometric convolutional neural network architecture. Input data included curvature and myelin values (vertex features) and the arrangement of vertices within the region of interest on the cortical surface. The output data is the vertex-wise polar angle/eccentricity values. Network parameters were optimized to reduce the difference between prediction and ground truth using a smooth L1 loss function.

Using geometric deep learning, we developed a neural network capable of predicting the detailed functional organization of the human visual hierarchy (Figure 2c). The visual hierarchy is comprised of a number of functionally specialized cortical visual areas[10], nearly all of which are organized retinotopically[11,12]. That is, the spatial organization of the retina (and hence the visual field) is maintained in each of these cortical visual areas - albeit in a distorted fashion due to properties such as cortical magnification (i.e., there are more neurons dedicated to processing foveal vs. peripheral information)[13,14]. Although distorted, this retinotopic mapping is known to be similar across individuals and the mapping functions between the visual field and the brain, at least in early visual cortex, can be described by relatively simple mathematical models[15,16]. However, considerable inter-individual variation does exist, and this

variation has been shown to be directly related to variability in cortical folding patterns and other anatomical features such as the degree of myelination[17–21]. It was our aim, therefore, to develop a neural network capable of learning the complex relationship between the functional organization of visual cortex and the underlying anatomy. To this end, we trained our network using the most comprehensive and detailed retinotopy dataset openly available – that from the Human Connectome Project (HCP)[22]. This dataset includes high field (7T) fMRI retinotopic mapping data of a large number of participants along with their anatomical data, serving as an ideal dataset for our endeavor. Although we show its utility for retinotopic mapping, geometric deep learning is flexible and well-suited for an abundance of other applications, both within the field of neuroscience[23–25] and beyond.

2. Methods

Participant information and data overview

Comprehensive information about the HCP 7T retinotopy dataset is given elsewhere[22]. Briefly, complete retinotopic mapping and structural data were acquired for 181 participants (109 females, age 22-35) following the HCP protocol[22,26]. All participants had normal or corrected-to-normal visual acuity. Structural image acquisition included T1w and T2w structural scans at 0.7mm isotropic resolution at a customized Siemens 3T Connectome scanner[26]. White and pial cortical surfaces were reconstructed from the structural scans using FreeSurfer (http://surfer.nmr.mgh.harvard.edu/) and were aligned across participants to the HCP 32k fs_LR standard surface space[27]. Myelin maps were determined by the ratio of T1w/T2w images[28] and appropriately normalized to account for B1+ transmit effects[27].

Whole-brain fMRI data were acquired using a 7T Magnetom scanner at a resolution of 1.6mm isotropic and 1s TR. The data were processed using the HCP pipeline[27], which involved correction for head motion and EPI spatial distortion, alignment of the fMRI data with the HCP standard surface space, and denoising for spatially specific structured noise. The data produced (in CIFTI format) by the pipeline consists of 91,282 grayordinates: 32,492 cortical vertices per hemisphere and 26,298 subcortical voxels with approximately 2mm spatial resolution. All data are publicly available on the ConnectomeDB database (https://db.humanconnectome.org).

Retinotopic mapping details

For our modeling, we used previously analyzed retinotopic mapping data[22] – made available at https://balsa.wustl.edu/study/show/9Zkk as scene files. These files can be visualized using Connectome Workbench and are efficiently handled using the MATLAB GIFTI library and Fieldtrip toolbox. For complete details on the experimental design and analysis procedure used to produce the maps, please refer to Benson et al. (2018)[22]. In brief, retinotopic stimuli were constrained to a circular region with a diameter of 16º. They consisted of slowly moving apertures of achromatic pink-noise background with dynamic colorful texture object, which was designed to elicit strong neural responses throughout the visual hierarchy – including in higher order visual areas. The experiment consisted of 6 runs with different set of apertures: wedges, rings and bars. Each aperture was presented twice, changing the direction of movement between runs, except for bar stimuli runs which were identical. Ring stimuli consisted of 8 cycles of a ring expanding away from the center or contracting towards the center with a period of 32s. Wedges stimuli consisted of 8 cycles of 90º wedges rotating across the visual field counterclockwise or clockwise with a period of 32s. Finally, each bar stimuli run consisted of bars with different orientations (4 orientations)

moving across different directions in visual field. During each condition, participants were to attend to a semitransparent dot located at the center of the display and report whenever the color of the dot changed (via a button press) to ensure their gaze were fixed at the center of the display.

Population receptive field (pRF) modeling[22,29], was then used to construct the retinotopic maps. Essentially, this modeling procedure estimates the spatial sensitivity profile within the visual field to which a grayordinate is responsive (i.e., its receptive field). For this, the fMRI time series elicited by the retinotopic mapping stimuli described above are modeled as the sum of a stimulus-related time series and a baseline time series. The stimulus-related response is then obtained by computing the dot product between the stimulus aperture time series and a 2D isotropic Gaussian (representing the pRF), applying a non-linearity component and rescaling the result by a gain factor, and then convolving it with a canonical hemodynamic response function. The modeled fMRI time series is optimized in order to best match the real fMRI time series by changing the 2D isotropic Gaussian parameters, such as its center position (x and y coordinates) and size, as well as the gain factor. Once the best parameters are determined, it is possible to construct retinotopic maps using each grayordinate's preferred location in the visual field (i.e., its pRF's center location). Polar angle maps reflect the polar angle in visual field to which a grayordinate is most responsive, while eccentricity maps reflect the distance from the center of the visual field (i.e., the fixation point) to which a grayordinate is most responsive.

Definition of the region of interest

As mentioned, the data files consisted of 91,282 grayordinates: 32,492 cortical surface vertices per hemisphere and 26,298 subcortical voxels. Here, we restricted the analysis to the cortical surface, hence ignoring the subcortical voxels. We further

restricted our region of interest (ROI) to only include vertices within visual cortex using a surface-based probabilistic atlas[7]. We modified the atlas slightly by extending the early visual cortex (V1, V2, V3) ROIs to include the foveal confluence[30] as well as combining ventral and dorsal components which are labelled separated in the original atlas (Figure 1). We also excluded the frontal eye fields (FEF) due to its discontinuity with the remaining areas. After selecting only vertices within this region of interest, we were left with 3,267 vertices for the left hemisphere (3,219 vertices of the right hemisphere), adding up to 19,024 connections (edges) between pairs of vertices (18,760 edges for the right hemisphere).

Data for deep learning

Our goal was to develop and train a deep neural network to predict the functional organization of the visual cortex (i.e., the retinotopic maps) from underlying anatomical properties (Figure 2). The anatomical data serving as input to the neural network included curvature and myelin values (both of which have been shown to be related to the functional organization of visual cortex) as well as the connectivity among the ROI surface vertices and their spatial disposition. The output of the model was either the polar angle or eccentricity value for each vertex of the cortical surface model. Note that polar angle values are cyclical (i.e., 360º = 0º). To avoid issues with the discontinuity in values, we shifted the polar angle values so that each hemisphere was trained with the point of wrap-around (from 360º to 0º) positioned at the horizontal meridian in the contralateral hemifield. After training, the values were shifted back to their original range.

For the modeling procedure, the data was represented on the midthickness cortical surface model (S1200_7T_Retinotopy181.L(R).midthickness_MSMAll.32k_fs_LR.surf.gii). Importantly, this model is a fiducial surface that maintains the actual geometry

of the brain. Our results are, however, displayed using a spherical surface model (S1200_7T_Retinotopy181.L(R).sphere.32k_fs_LR.surf.gii) for visualization purposes. This permits all visual areas to be displayed in a single view.

Developing the neural network

Developing the neural network involved three main steps: (1) *training* the neural network, (2) hyperparameter *tuning*, and (3) *testing* the model. Prior to the training step, the 181 participants from the HCP dataset were randomly separated into three independent datasets: training (161 participants), development (10 participants) and test (10 participants) datasets. These datasets were used in each of the above three steps, respectively. During *training*, the network learned the correspondence between the retinotopic maps and the anatomical features by exposing the network to each example in the training dataset. Hence, the empirically-derived retinotopic maps served as the “ground truth” and the parameters of the neural network were optimized to minimize the difference between the predicted and empirical retinotopic maps. Model hyperparameters, such as the number of layers, were *tuned* (i.e., optimized) by inspecting model performance using the development dataset. Finally, once the final model was selected, the network was *tested* by assessing the predicted maps for each participant in the test dataset. This procedure was followed for each hemisphere and each type of retinotopic mapping data (i.e., polar angle and eccentricity) separately, resulting in four predictive models.

Model architecture and training

We implemented a spline-based convolutional neural network (SplineCNN)[31] on cortical surfaces using PyTorch Geometric[32], a geometric deep learning extension of PyTorch. Among the broad range of methods defining convolution operations on

irregular structured data, such as graphs and surfaces, spline-based convolution is amidst the spatial filtering approaches. In these approaches, filters aggregate information locally, around a vertex's neighborhood. These filters exploit information given by relative positions of the neighbor vertices with respect to the reference vertex in addition to the information encoded in the connectivity, edge weights, and vertex features. Thus, SplineCNN aggregates vertex features in local neighborhoods weighted by learnable parameters of a continuous kernel function[31].

Following the original study, we denote spline-based convolutional layers as SConv(k,M_{in},M_{out}) of which k is the kernel size, M_{in} is the number of input feature maps, and M_{out} is the number of output feature maps (Figure 2). Our final model architecture included 12 convolutional layers with the number of feature maps increasing to a maximum of 32 feature maps per layer, and then decreasing back to one final feature map. A non-linear activation function, the exponential linear unit (ELU), was used after each convolution layer. Batch normalization and dropout (p=0.10) were applied to all layers, except the last. In our model, we only used Cartesian coordinates to compute relative distance between vertices and the degree of B-spline basis m = 1, due to the fact that models with those parameters showed better performance than the models using higher degree of B-spline basis or using spherical coordinates in previous experiments[31]. Training was carried out for 200 epochs with a batch size of 1 and a learning rate at 0.01 for 100 epochs that was then adjusted to 0.005, using Adam optimizer. Our models' learning objective was to reduce the difference between predicted retinotopic map and ground truth (i.e., the empirical retinotopic map). This mapping objective is measured by the smooth L1 loss function (Equation 1).

$$\boldsymbol{loss}(\hat{\boldsymbol{y}}^*, \boldsymbol{y}^*) = \frac{\boldsymbol{1}}{\boldsymbol{n}} \sum_{\boldsymbol{i}} \boldsymbol{z_i} \qquad \textbf{(1)}$$

where y^* and $\hat{y}^*$ are, respectively, the empirical and predicted value weighted by the

explained variance (R^2) from the pRF modeling procedure[22], n is the total number of vertices, and z_i is given by:

$$z_i = \begin{cases} 0.5(\hat{y}^*{}_i - y^*{}_i)^2, if\ |\hat{y}^*{}_i - y^*{}_i| < 1 \\ |\hat{y}^*{}_i - y^*{}_i| - 0.5, \qquad otherwise \end{cases}$$

Models were implemented using Python 3.7.3 and Pytorch 1.2.0. Training was performed on a high-performance computing cluster using NVIDIA Tesla V100 Accelerator units.

Hyperparameter tuning

Hyperparameter tuning was conducted by considering the performance of the model on predicting retinotopic maps of participants within the development dataset. We used the mean absolute error to compute the difference between predicted and expected values (ground truth). During the training stage, we monitored the decrease of the loss function (i.e., the prediction error on the training dataset) and the decrease of the mean absolute error (i.e., the prediction error on the development dataset). We also visually monitored predicted retinotopic maps using the development dataset, as we observed that despite the prediction error within the development dataset reaching a plateau around 100 epochs, some of the details in the predicted retinotopic maps continued to vary with additional training. To allow for some of this additional fine tuning, we fixed our training to 200 epochs.

To find the most appropriate number of layers, we trained five different models per number of layers, varying from 1 to 20 layers, as the performance of the final model fluctuated due to random initialization of the learnable parameters. We selected the number of layers based on the performance of the trained network, which was estimated by the least difference between two angles, given by:

$$\mathrm{MIN}(|\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}|, |\hat{\boldsymbol{\theta}} - \boldsymbol{\theta} + \mathbf{2}\boldsymbol{\pi}|, |\hat{\boldsymbol{\theta}} - \boldsymbol{\theta} - \mathbf{2}\boldsymbol{\pi}|) \quad \textbf{(2)}$$

for $0 < \theta < 2\pi$. To this end, we considered two factors: (1) how well model prediction matched the ground truth (error) and (2) how much a predicted map differed from the others (individual variability). To estimate the error, we computed the smallest difference between the predicted and the empirical (ground truth) polar angle values in a vertex-wise manner and averaged across all participants in the development dataset and within the various ROIs. The other factor, individual variability, was determined by the smallest difference between a specific predicted map and each other predicted map in the development dataset in a vertex-wise manner and averaged across all combinations (there are 10 participants in the development dataset, hence 9 combinations), and then average across participants and within the various ROIs. This was performed for all models (5 models per number of layers) and we selected the number of layers that resulted in the best performance across all the ROIs in both metrics (12 layers; see Supplementary Figure 3). This selection was based on mapping polar angle values in the left hemisphere; however, we tested a range of layers using the right hemisphere data and found similar results. Finally, once the appropriate number of layers was established, we chose the best model among the 5 models that were trained with the same hyperparameters but with different learnable parameters (due to the random initialization) to evaluate on the test dataset. For the eccentricity models (left and right hemispheres), we kept the same hyperparameters as optimized for polar angle models.

Model testing and evaluation metric

The selected models were evaluated based on their performance on the test dataset considering the same two factors, error and individual variability. Importantly, the test dataset was never seen by the network nor by the researchers until the deep learning

architecture was finalized and the final model selected. Again, the error was determined by the smallest difference between the predicted and the empirical polar angle values in a vertex-wise manner and averaged across all participants – this time in the test dataset. Individual variability was determined by the smallest difference between a specific predicted map and each other predicted map in the test dataset in a vertex-wise manner and averaged across all combinations (9 combinations), and then averaged across participants.

To evaluate the importance of the spatial organization of the anatomical features and their variability for the prediction of individual differences in retinotopic maps, the effects of modifying the input data in two ways was tested: (1) randomly shuffling the anatomical features and (2) setting the input features equal to a constant value. By randomly shuffling the features across the ROI vertices for each participant in the test dataset, the spatial organization of the anatomical information was disrupted. If the spatial disposition of anatomical features on the cortical surface is indeed important for predicting individual variability in retinotopic maps, it would then be expected that the model would generate less accurate retinotopic maps (i.e., higher error between the predicted and empirical maps and less variable across individuals) compared to those predicted when using intact anatomical features. Next, to test the effect of completely removing the variability present in the input features, all curvature and myelin values of each participant in the test dataset was set equal to the mean curvature (0.027) and the mean myelin (1.439) values, respectively. Note that the mean values were computed using all participants from the training dataset. By removing the variability of anatomical features, the model was expected to generate the same retinotopic map for all participants (i.e., no individual variability). Hence without the ability to vary across participants, less accurate predictions (i.e., greater

error measures) were expected compared to using the intact (and variable) features.

Data and code availability

The data used in this study in publicly available at BALSA (https://balsa.wustl.edu/study/show/9Zkk). The pre-trained models with accompanying Python source code will be available upon publication.

3. Results and Discussion

Retinotopic mapping with deep learning

Retinotopic mapping typically involves empirical data collection performed by showing visual stimuli, which are temporally and spatially controlled, to a participant lying in a MRI scanner[33,34]. Analysis of the fMRI signals elicited by such stimuli allows the visual field location encoded by each responsive voxel to be estimated. The visual field locations are usually defined in polar coordinates resulting in two retinotopic maps: one representing the polar angle (or clock position) and the other representing the eccentricity (or distance away from fixation). An example of these empirically derived retinotopic maps for a single individual is shown in Figure 3 (top row). From these, it can be seen that the polar angle and eccentricity maps are organized roughly orthogonal to one another – most easily discerned in early visual cortex (V1, V2, and V3). Since the human visual system is comprised of multiple retinotopically organized cortical areas, these maps are also used to demarcate the boundaries of the various areas. The polar angle maps are particularly useful for this as the boundaries between areas can be identified by reversals in the progression of the polar angle values. In Figure 3, for example, the superior boundary between V1 and V2 is found by locating the band at which the polar angle reverses from progressing from red to yellow, back

to red. Likewise the inferior boundary between V1 and V2 is found where the polar angle representation reverses from progressing from red to purple, back to red.

To map retinotopically organized areas without the need to collect functional data, we built a geometric convolutional neural network (Figure 2c) able to predict the polar angle and eccentricity maps from anatomical properties alone (i.e., curvature and myelin values). Figure 3 illustrates the predicted retinotopic maps with their correspondent ground truth (i.e., empirically measured retinotopic maps) for a single participant in the test dataset. Qualitatively, the predicted maps bear a remarkable resemblance to the ground truth, appearing to accurately predict the main features of both polar angle and eccentricity maps. To aid comparison between the empirical and predicted maps a grid of isopolar angle and isoeccentiricty lines has been overlaid upon the maps in early visual cortex. The grid was drawn based on the ground truth data and then positioned identically on the predicted maps. Note how well the visual area boundaries (i.e., the isopolar angle contours) match the predicted maps, running directly through the polar angle reversals. A similar quality of fit can be seen for the isoeccentricty lines. Although the grid is focused on early visual areas (due to their organization being most well understood and hence most reliable to outline), it is important to note that the neural networks are able to predict the retinotopic organization far beyond early visual cortex, throughout the visual hierarchy. Being smaller in size, it is difficult to see the finer retinotopic details in these areas, however, the predicted maps in these areas do appear to reflect the general organizational features present in the empirical data – such as additional polar angle reversals throughout the hierarchy (i.e., ventral, lateral, and dorsal to early visual cortex).

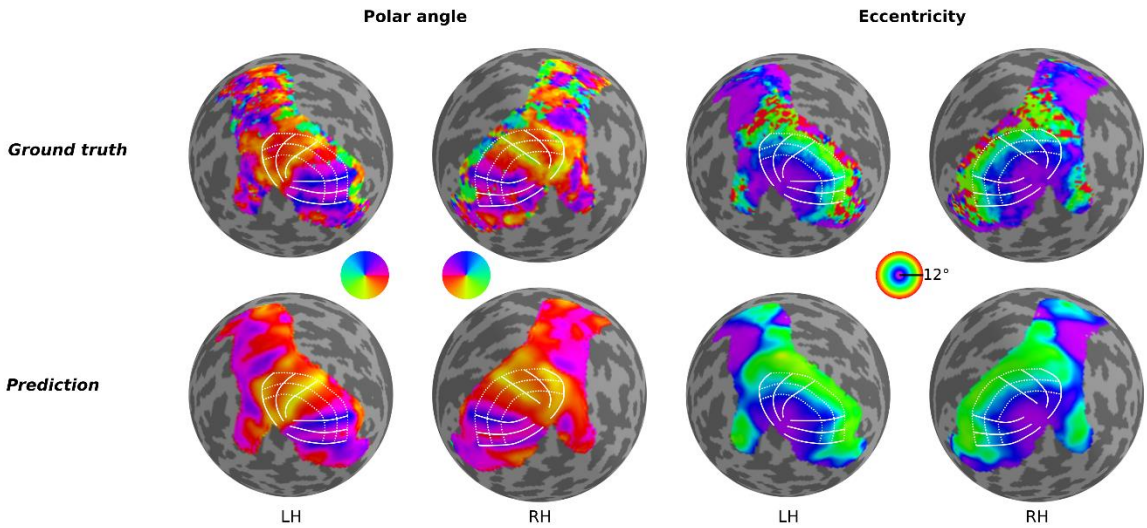


Figure 3 - Retinotopic mapping with geometric deep learning. Upper row shows empirical (ground truth) polar angle and eccentricity maps for both left (LH) and right (RH) hemispheres of a single individual from the test dataset (Participant 1). Lower row shows the predicted polar angle and eccentricity maps for each hemisphere of the same participant. Polar angles vary from 0º to 360º, while eccentricity values varied from 0º to 12º as indicated by the color wheels. The white grid illustrates isoeccentricty (dashed) and isopolar angle (solid) lines associated with early visual cortex (V1, V2, and V3). Note that V2 and V3 are split around V1 and are comprised of dorsal and ventral components.

Individual variability in predicted maps

It is clear that the geometric deep learning models can predict the basic retinotopic organization of human visual cortex and that the predicted maps appear to be in line with the empirically measured maps for one participant, but are these models capable of predicting individual variability across participants? Figure 4 shows the results for 4 additional participants from the test dataset. For brevity, we only show the polar angle maps from the left hemisphere, but all the predicted maps for each test dataset participant are included as supplementary material (Supplementary Figures 1 and 2). Additionally, the mean prediction error of each of the four predictive models are summarized in Supplementary Table 2. In Figure 4a, the empirical polar angle maps are marked by unusual and/or discontinuous polar angle reversals. In the first three maps (Figure 4a, Participants 2-4), a discontinuous representation of the lower vertical meridian (that is, the boundaries between V1 and dorsal V2 as well as between dorsal

V2 and dorsal V3) can be observed – indicated by the gray lines. Importantly, these unique variations were correctly predicted by our model. Perhaps even more striking, we see that these borders have merged to form a Y-shape for Participant 5, which was also observed in the predicted map.

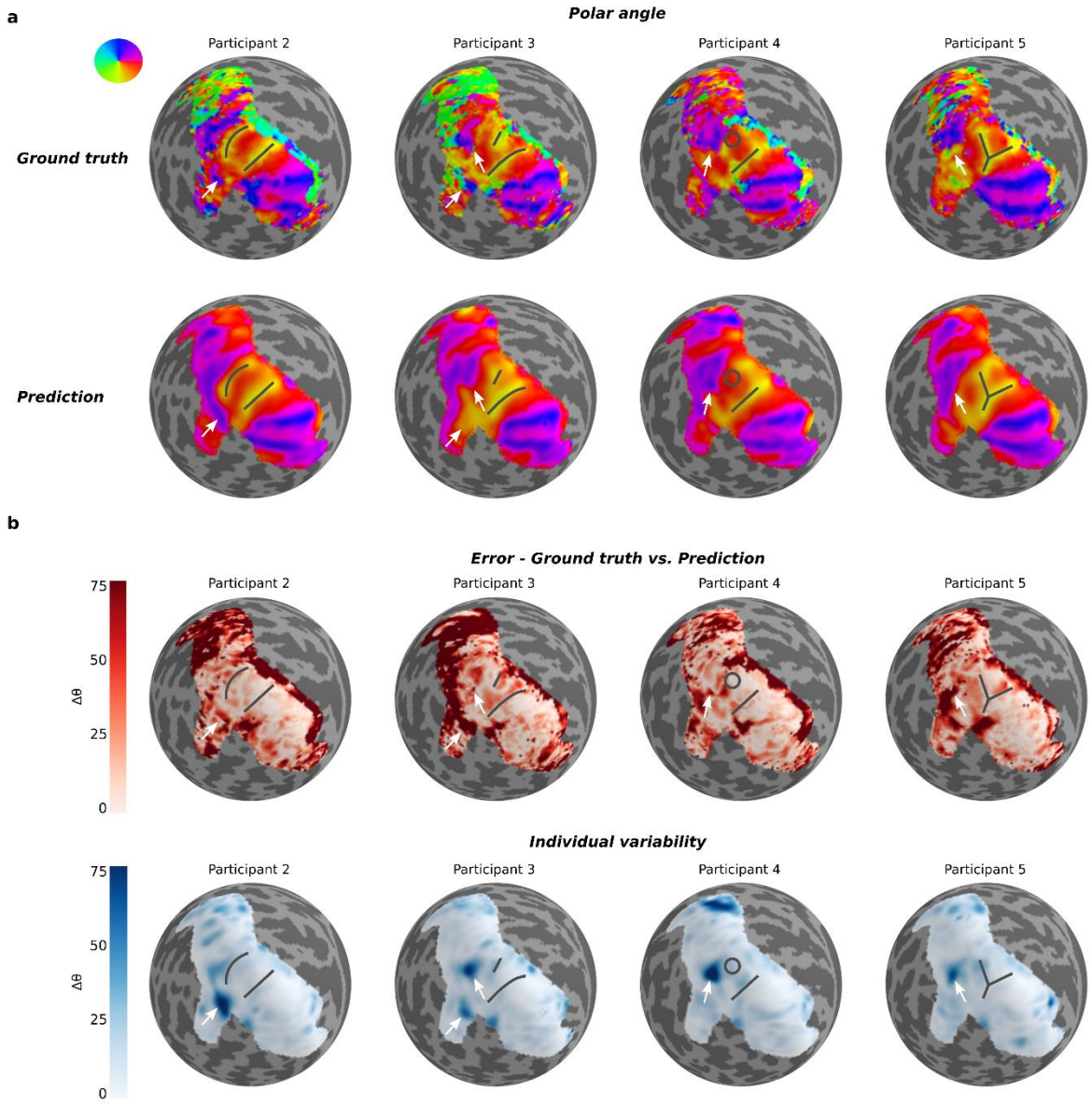


Figure 4 - Individual variability in retinotopic maps. a, Upper row shows empirical (ground truth) polar angle maps and lower row shows their respective predicted polar angle maps for four participants in the test dataset. Grey lines indicate unique polar angle representations that were predicted by our model. **b**, Upper row shows the prediction error of each individual map and lower row shows how variable the same given map is from all the other predicted maps in the test dataset. White arrows indicate regions that differed the most among predicted maps.

To further assess the performance of our models, the smallest difference between the predicted and the empirical (ground truth) polar angle values was determined in a vertex-wise manner (upper row, Figure 4b). This error estimate was then displayed on the cortical surface meshes to inspect where our models perform well and where they suffer. Predictions differed most from the ground truth near the foveal confluence (i.e., where the V1, V2, and V3 foveal representations merge), at the most peripheral eccentricities, and in some of the higher order visual areas (particularly those located dorsally in the intraparietal sulcus). Notably, each of these regions are known to be problematic when measuring retinotopic maps empirically. For example, it has been shown that retinotopic measurements within the foveal confluence are often not well represented due to instrumental and measurement limitations[35]. The errors seen at the most peripheral eccentricities likely manifest from the suppressive effects seen in the cortex for retinotopic locations immediately outside the most peripheral limit of the visual stimuli[36,37]. Finally, the higher order areas are also known to be more difficult to map empirically using conventional procedures due to their smaller size and their higher degree of selectivity to specific types of stimuli[35]. Together, this brings us to an important caveat: the empirical maps, although treated as ground truth, may themselves misrepresent the underlying retinotopic organization. Accordingly, Figure 5 shows the average explained variance (R^2) associated with the original pRF analysis of the empirical retinotopic mapping data[22] and as such, can be seen as a proxy for the quality of our ground truth. Pertinently, it shows that quality of the empirical data is indeed lowest in the regions where our model predictions were most error-prone – e.g., in the foveal confluence (Figure 5, arrow a), at the most peripheral eccentricities (Figure 5, arrow b), and in some of the higher-order areas (Figure 5, arrow c).

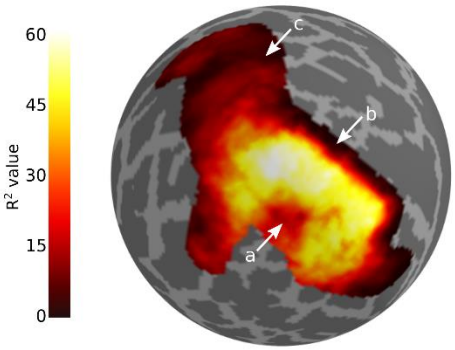


Figure 5 – Quality of the empirical retinotopic mapping data. Average explained variance (R^2) map of pRF modeling[22] from all participants in the test dataset.

Similar to the error estimates, we were interested in quantifying the extent to which each individual's predicted map within the test dataset varied from the others. For this we computed an individual variability estimate as the smallest difference between a particular predicted map and each of the other predicted maps in the test dataset in a vertex-wise manner and averaged across all comparisons. Results are shown in Figure 4b, lower row. As expected, locations that more prominently differed among predicted maps in early visual cortex were the areas in and near the unique discontinuities seen around the lower vertical meridian lines in the polar angle maps (Figure 4, white arrows).

Variability guided by anatomical features

To demonstrate the importance of the anatomical features in shaping individual differences in the functional organization of the visual cortex, we predicted two new sets of retinotopic maps for each participant after (1) randomly shuffling the anatomical input features and (2) setting the input features equal to a constant value. We expected that disrupting the input features in these ways would result in both higher error and lower individual variability across predicted maps. Figure 6a (upper row) shows the average error and the individual variability estimates across all participants. The

highest errors were again present in the foveal confluence, higher-order visual areas, and near the outermost eccentricities. These errors, however, can be seen when using the actual empirical values as input as well as the disrupted values, and likely reflect the lack of reliable training data in these regions (Figure 5). Contrasting the three sets of maps, however, reveals that the errors were generally higher throughout visual cortex when the input values were either shuffled or constant when compared to intact input values, indicating that the spatial organization of the anatomical features is being used by the network for accurate predictions. Interestingly, disrupting the anatomical input features particularly increases the error in the dorsal regions of early visual cortex (Figure 6a), which is where most the variability was seen across individuals in Figure 4. Complementary to this, we also see that the proper spatial organization of the anatomical features is important for generating individual variability in our model predictions (Figure 6a, lower row). The variability across predicted maps decreased when the input values were randomly shuffled. Moreover, as expected the variability between predicted maps was equal to zero when the input values were the same across all vertices; without variability in the input features, the model will always predict the same maps.

To summarize the vertex-wise error and individual variability estimates in Figure 6a, we computed the mean error and the mean individual variability across the visual cortex (Figure 6b). For this, we first grouped the 21 visual areas (Figure 1) into six different clusters (i.e., anatomically contiguous areas with similar positions in the hierarchy), as indicated in Figure 6b. Disrupting the spatial organization of input features led to both an increased mean error between predicted maps and ground truth as well as decreased mean individual variability across all clusters. Together,

these results show that the anatomical features were critical for predicting the functional organization of the visual cortex at the individual level.

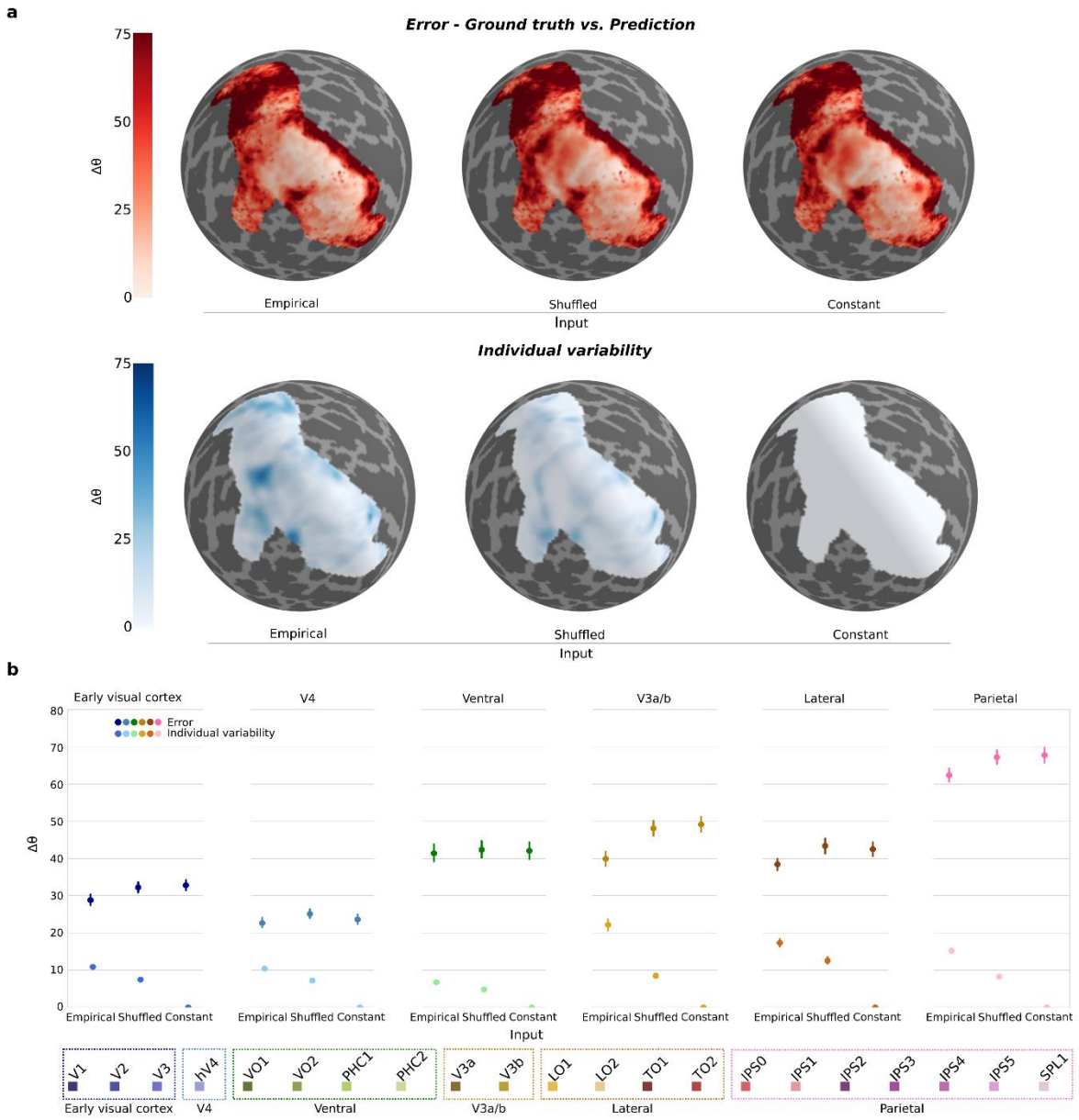


Figure 6 - Individual variability in predicted maps is determined by anatomical features. a, Upper row shows the prediction error across participants in the test dataset using the empirical input features, shuffled input features, and constant input features. Lower row shows how variable predicted maps are from all the other predicted maps in the test dataset on average for each input condition. **b**, Mean and bootstrapped (n=1,000) 95% confidence interval of the average error and individual variability of predicted maps for each input condition within different visual areas clusters.

4. Conclusion

This study shows that deep learning can be used to predict the detailed functional organization of visual cortex from anatomical features alone by building neural networks able to predict retinotopic maps throughout the visual cortical hierarchy. Remarkably, the networks were able to predict unique features in these maps showing sensitivity to individual differences. Previous work has been done to map the correspondence between brain anatomy and retinotopic maps in individuals by warping an atlas or template[17–19] according to their specific anatomical data. While these approaches are able to provide reasonable estimates of their retinotopic maps, they have not – when using anatomical information alone – been able to capture the detailed idiosyncrasies seen in the actual measured maps of those individuals. By not assuming a prior structure-function relationship our deep learning framework provides a more flexible approach, able to produce more specific retinotopic maps.

One key aspect of this work was the use of a geometric approach, which permitted deep learning to be performed directly on the cortical surface model. Doing so, we were able to take advantage of the actual structure of the brain and hence more appropriately represent the data to the neural network. It is not rare for data to better be represented in a non-Euclidean space than in the Euclidean one. For example, also within the field of neuroscience, brain functional networks are often represented as graphs (i.e., connectomes) in which connections between two brain regions do not necessarily represent the physical distance between them. Geometric deep learning, being also well-suited for graph data[6], offers the field a powerful tool given the growing interest in predicting aspects of human behavior from these functional connectivity graphs. Beyond neuroscience lies a wide range of other potential applications for these geometric deep learning techniques – including application to topological data

in earth science[40], genetic and protein structure data in biological science[41–43], drug discovery in chemistry science[44], cosmological data in physical science[45], and more.

Acknowledgments

This work was supported by the Australian Research Council (DE180100433).

Author contributions

F.L.R., S.B., and A.M.P. conceptualized the study. F.L.R. and A.M.P. designed the study and analyses. F.L.R. carried out the modeling procedure and analyses. F.L.R. and A.M.P. wrote the manuscript. All authors reviewed the manuscript.

Competing financial interests

The authors declare no competing financial interests.

References

1. Fukushima, K. Neocognition: A Self-organizing Neural Network Model for a Mechanism of Pattern Recognition Unaffected by Shift in Position. *Biol. Cybern.* **36**, 193–202 (1980).
2. Lecun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436–444 (2015).
3. Hubel, D. H. & Wiesel, T. N. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol.* **160**, 106–154 (1962).
4. DJ, F. & DC, V. E. Distributed hierarchical processing in the primate cerebral cortex. *Cereb. Cortex* **1**, 1–47 (1991).
5. Kanwisher, N., McDermott, J. & Chun, M. M. The Fusiform Face Area: A Module in Human Extrastriate Cortex Specialized for Face Perception. *J. Neurosci.* **17**, 4302–4311 (1997).
6. Bronstein, M. M., Bruna, J., Lecun, Y., Szlam, A. & Vandergheynst, P. Geometric Deep Learning: Going beyond Euclidean data. *IEEE Signal Process. Mag.* **34**, 18–42 (2017).
7. Wang, L., Mruczek, R. E. B., Arcaro, M. J. & Kastner, S. Probabilistic maps of visual topography in human cortex. *Cereb. Cortex* **25**, 3911–3931 (2015).

8. Van Essen, D. C., Drury, H. A., Joshi, S. & Miller, M. I. Functional and structural mapping of human cerebral cortex: solutions are in the surfaces. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 788–795 (1998).

9. Dale, A. M., Fischl, B. & Sereno, M. I. Cortical surface-based analysis: I. Segmentation and surface reconstruction. *Neuroimage* **9**, 179–194 (1999).

10. Zeki, S. *et al.* A direct demonstration of functional specialization in human visual cortex. *J. Neurosci.* **11**, 641–649 (1991).

11. Grill-Spector, K. & Malach, R. The Human Visual Cortex. *Annu. Rev. Neurosci.* **27**, 649–677 (2004).

12. Van Essen, D. C. Organization of visual areas in macaque and human cerebral cortex. in *In The Visual Neuroscience* (eds. Chapula, L. & Werner, J.) 507–521 (Cambridge: MIT, 2004).

13. Cohen, R. A. Cortical Magnification. in *Encyclopedia of Clinical Neuropsychology* (eds. Kreutzer, J. S., DeLuca, J. & Caplan, B.) 718–719 (Springer New York, 2011). doi:10.1007/978-0-387-79948-3_1355

14. Cowey, A. & Rolls, E. T. Human cortical magnification factor and its relation to visual acuity. *Exp. Brain Res.* **21**, 447–454 (1974).

15. Schira, M. M., Wade, A. R. & Tyler, C. W. Two-Dimensional Mapping of the Central and Parafoveal Visual Field to Human Visual Cortex. *J. Neurophysiol.* **97**, 4284–4295 (2007).

16. Balasubramanian, M., Polimeni, J. & Schwartz, E. L. The V1-V2-V3 complex: quasiconformal dipole maps in primate striate and extra-striate cortex. *Neural Networks* **15**, 1157–1163 (2002).

17. Benson, N. C. *et al.* The retinotopic organization of striate cortex is well predicted by surface topology. *Curr. Biol.* **22**, 2081–2085 (2012).

18. Benson, N. C., Butt, O. H., Brainard, D. H. & Aguirre, G. K. Correction of Distortion in Flattened Representations of the Cortical Surface Allows Prediction of V1-V3 Functional Organization from Anatomy. *PLoS Comput. Biol.* **10**, (2014).

19. Benson, N. C. & Winawer, J. Bayesian analysis of retinotopic maps. *Elife* **7**, 1–29 (2018).

20. Hinds, O. P. *et al.* Accurate prediction of V1 location from cortical folds in a surface coordinate system. *Neuroimage* **39**, 1585–1599 (2009).

21. Abdollahi, R. O. *et al.* Correspondences between retinotopic areas and myelin maps in human visual cortex. *Neuroimage* **99**, 509–524 (2014).

22. Benson, N. C. *et al.* The Human Connectome Project 7 Tesla retinotopy dataset : Description and population receptive field analysis Essen. *J. Vis.* **18**, 1–22 (2018).

23. Seong, S.-B., Pae, C. & Park, H.-J. Geometric Convolutional Neural Network for Analyzing Surface-Based Neuroimaging Data. *Front. Neuroinform.* **12**, 1–13 (2018).

24. Zhao, F. *et al.* Spherical U-Net on Cortical Surfaces: Methods and Applications. 1–12 (2019).

25. Gopinath, K., Desrosiers, C. & Lombaert, H. Graph Convolutions on Spectral Embeddings for Cortical Surface Parcellation. *Med. Image Anal.* **54**, 297–305 (2018).

26. Van Essen, D. C. *et al.* The WU-Minn Human Connectome Project: An overview. *Neuroimage* **80**, 62–79 (2013).

27. Glasser, M. F. *et al.* The minimal preprocessing pipelines for the Human Connectome Project. *Neuroimage* **80**, 105–124 (2013).

28. Glasser, M. F. & Van Essen, D. C. Mapping human cortical areas in vivo based on myelin content as revealed by T1- and T2-weighted MRI. *J. Neurosci.* **31**, 11597–11616 (2011).

29. Kay, K. N., Winawer, J., Mezer, A. & Wandell, B. A. Compressive spatial summation in human visual cortex. *J. Neurophysiol.* **110**, 481–494 (2013).

30. Schira, M. M., Tyler, C. W., Breakspear, M. & Spehar, B. The foveal confluence in human visual cortex. *J. Neurosci.* **29**, 9050–9058 (2009).

31. Fey, M., Lenssen, J. E., Weichert, F. & Muller, H. SplineCNN: Fast Geometric Deep Learning with Continuous B-Spline Kernels. *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.* 869–877 (2018). doi:10.1109/CVPR.2018.00097

32. Fey, M. & Lenssen, J. E. Fast Graph Representation Learning with PyTorch Geometric. *Prepr. http//arxiv.org/abs/1903.02428* 1–9 (2019).

33. DeYoe, E. A. *et al.* Mapping striate and extrastriate visual areas in human cerebral cortex. *Proc. Natl. Acad. Sci.* **93**, 2382–2386 (1996).

34. Dumoulin, S. O. & Wandell, B. A. Population receptive field estimates in human visual cortex. *Neuroimage* **39**, 647–660 (2008).

35. Wandell, B. A. & Winawer, J. Imaging retinotopic maps in the human brain. *Vision Res.* **51**, 718–737 (2010).

36. Puckett, A. M., Mathis, J. R. & Deyoe, E. A. An investigation of positive and inverted hemodynamic response functions across multiple visual areas. *Hum. Brain Mapp.* **35**, 5550–5564 (2014).

37. Shmuel, A. *et al.* Sustained negative BOLD, blood flow and oxygen consumption response and its coupling to the positive response in the human brain. *Neuron* **36**, 1195–1210 (2002).

38. Griffis, J. C. *et al.* Retinotopic patterns of functional connectivity between V1 and large-scale brain networks during resting fixation. *Neuroimage* **176**, 1071–1083 (2017).

39. Park, B. Y., Tark, K. J., Shim, W. M. & Park, H. Functional connectivity based parcellation of early visual cortices. *Hum. Brain Mapp.* **39**, 1380–1390 (2018).

40. Muszynski, G., Kashinath, K., Kurlin, V., Wehner, M. & Prabhat. Topological data analysis and machine learning for recognizing atmospheric river patterns

in large climate datasets. *Geosci. Model Dev.* **12**, 613–625 (2019).

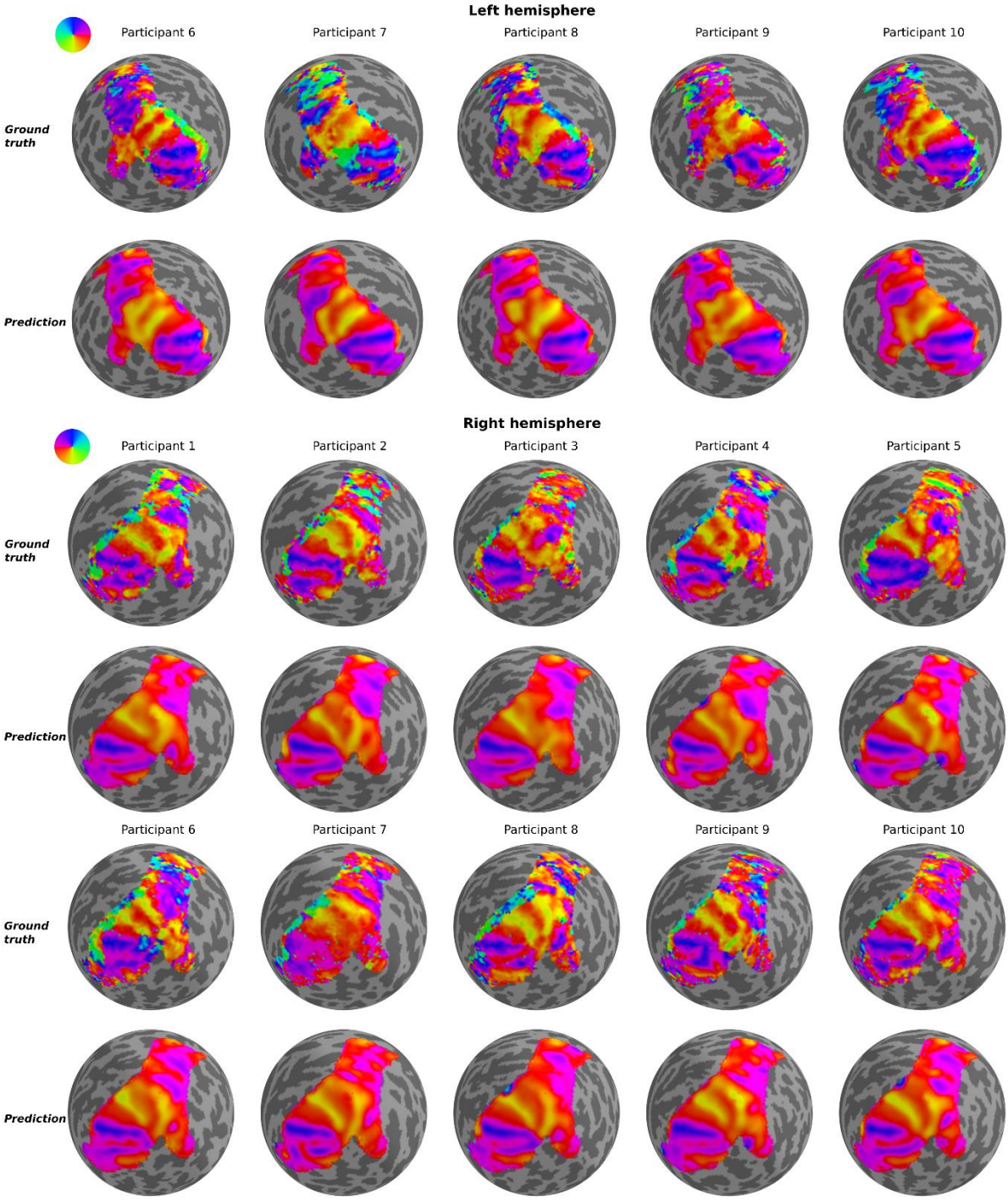
41. Gainza, P. *et al.* Deciphering interaction fingerprints from protein molecular surfaces. *bioRxiv* (2019). doi:10.1101/606202

42. Rossi, E., Monti, F., Bronstein, M. & Liò, P. ncRNA Classification with Graph Convolutional Networks. *ArXiv* (2019).

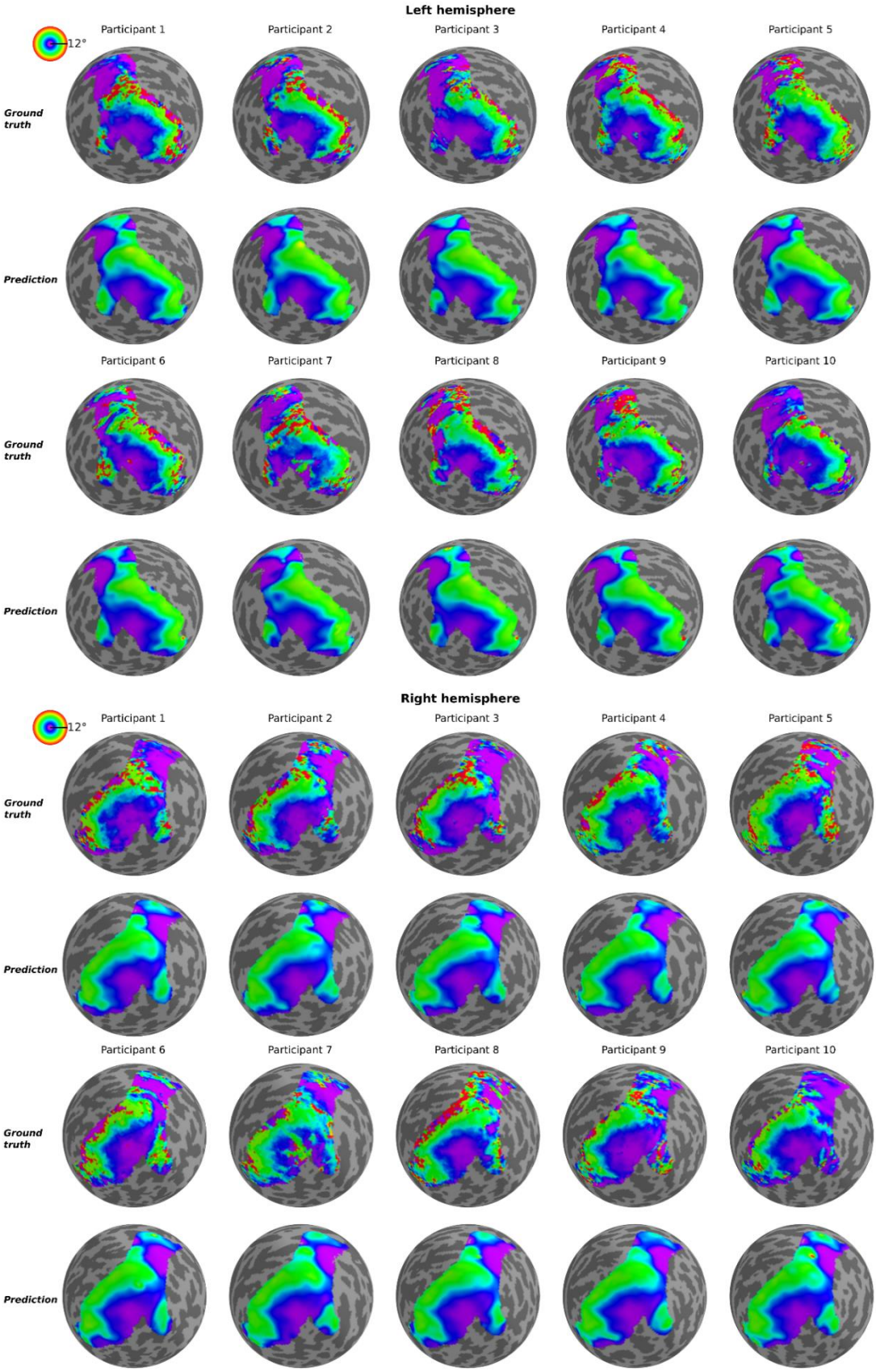
43. Eraslan, G., Avsec, Ž., Gagneur, J. & Theis, F. J. Deep learning: new computational modelling techniques for genomics. *Nat. Rev. Genet.* (2019). doi:10.1038/s41576-019-0122-6

44. Hop, P., Allgood, B. & Yu, J. Geometric Deep Learning Autonomously Learns Chemical Features That Outperform Those Engineered by Domain Experts. *Mol. Pharm.* **15**, 4371–4377 (2018).

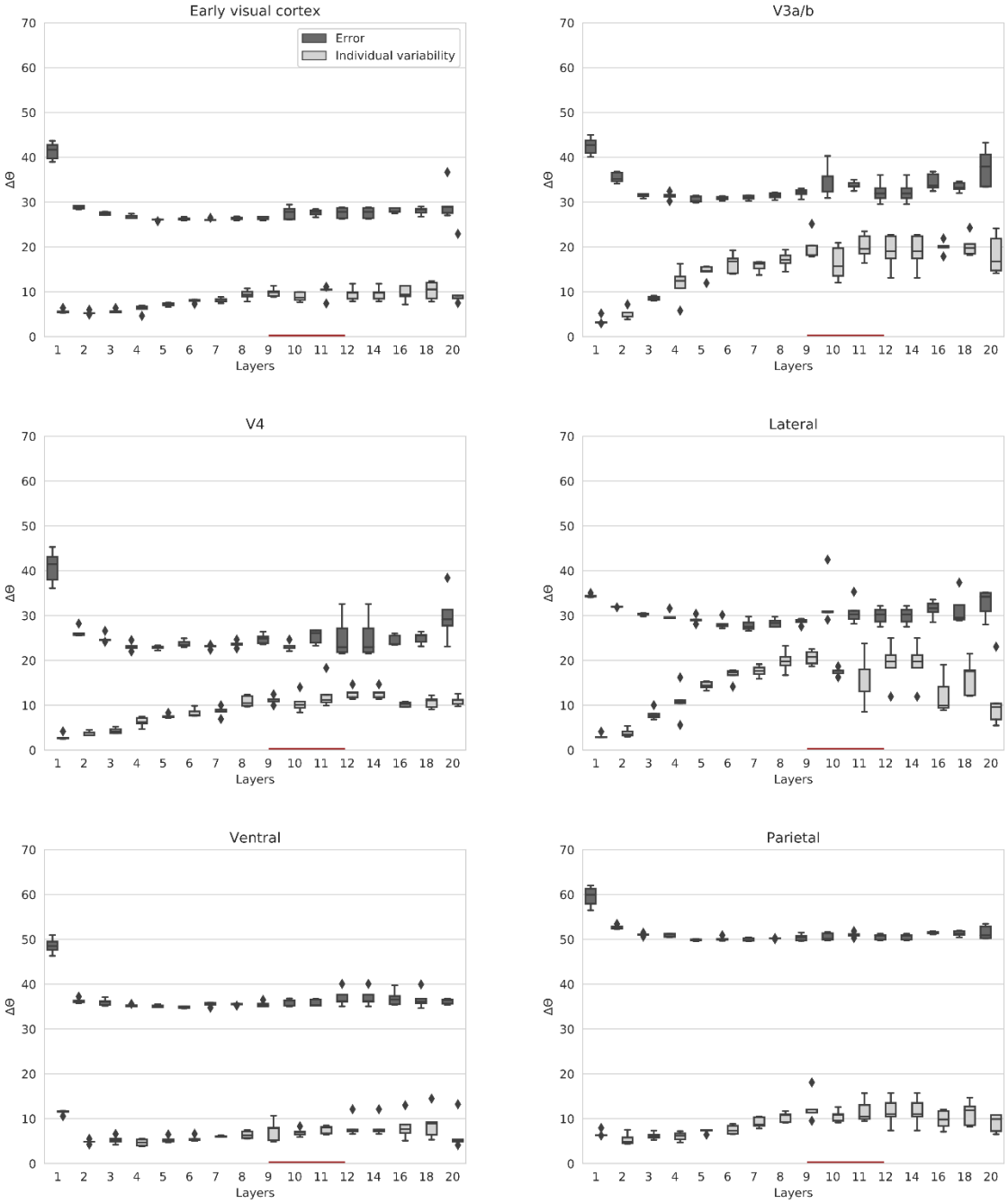
45. Perraudin, N., Defferrard, M., Kacprzak, T. & Sgier, R. DeepSphere: Efficient spherical convolutional neural network with HEALPix sampling for cosmological applications. *Astron. Comput.* **27**, 130–146 (2019).



Supplementary Figure 1 – Polar angle mapping with geometric deep learning. Upper panel shows empirical (ground truth) and predicted polar angle maps for the left hemisphere of the remaining individuals in the test dataset (polar angle maps of participants 1-5 are in the main manuscript). Lower panel shows empirical and predicted polar angle maps for the right hemisphere of all individuals in the test dataset. Polar angles vary from 0º to 360º as indicated by the color wheels.



Supplementary Figure 2 - Eccentricity mapping with geometric deep learning. Upper panel shows empirical (ground truth) and predicted eccentricity maps for the left hemisphere of all individuals in the test dataset. Lower panel shows empirical and predicted polar angle maps for the right hemisphere of all individuals in the test dataset. Eccentricity varied from 0º to 12º as indicated by the color wheels.



Supplementary Figure 3 – Model performance with varying number of layers. Box plots represent the mean prediction error (dark gray) and mean individual variability (light gray) of 5 different models trained for polar angle mapping in the left hemisphere – separated by visual area cluster. Evaluation of the performance of the models was conducted using the development dataset, as part of the hyperparameter tuning step. Models varied in the number of convolutional layers. Note that the mean error reduces as models get deeper (i.e., more layers) while the individual variability increases. However, performance starts to deteriorate for models with more than 12 layers. Models with 9 to 12 layers showed the best performance across all visual areas clusters.

Supplementary Table 1 - Participants' HCP identification number. Identification number of the participants within the test and development datasets.

DATASET	HCP ID	PARTICIPANT NUMBER
TEST	680957	1
	191841	2
	617748	3
	725751	4
	198653	5
	191336	6
	572045	7
	601127	8
	644246	9
	157336	10
DEVELOPMENT	186949	-
	169747	-
	826353	-
	825048	-
	671855	-
	751550	-
	318637	-
	131722	-
	137128	-
	706040	-

Supplementary Table 2 - Mean prediction error of the four predictive models on the test dataset. The mean of the average vertex-wise error (smallest difference between the predicted and empirical values) of vertices within the early visual cortex (V1, V2 and V3), as well as in all 21 visual areas, of each model on the test dataset.

POLAR ANGLE		
		Mean Δθ ± std
LEFT	Early visual cortex	28.85 ± 24.98
	All	41.10 ± 26.32
RIGHT	Early visual cortex	29.18 ± 26.26
	All	35.14 ± 21.71
ECCENTRICITY		
		Mean Δθ ± std
LEFT	Early visual cortex	1.15 ± 1.12
	All	1.83 ± 1.37
RIGHT	Early visual cortex	1.08 ± 1.11
	All	1.69 ± 1.24