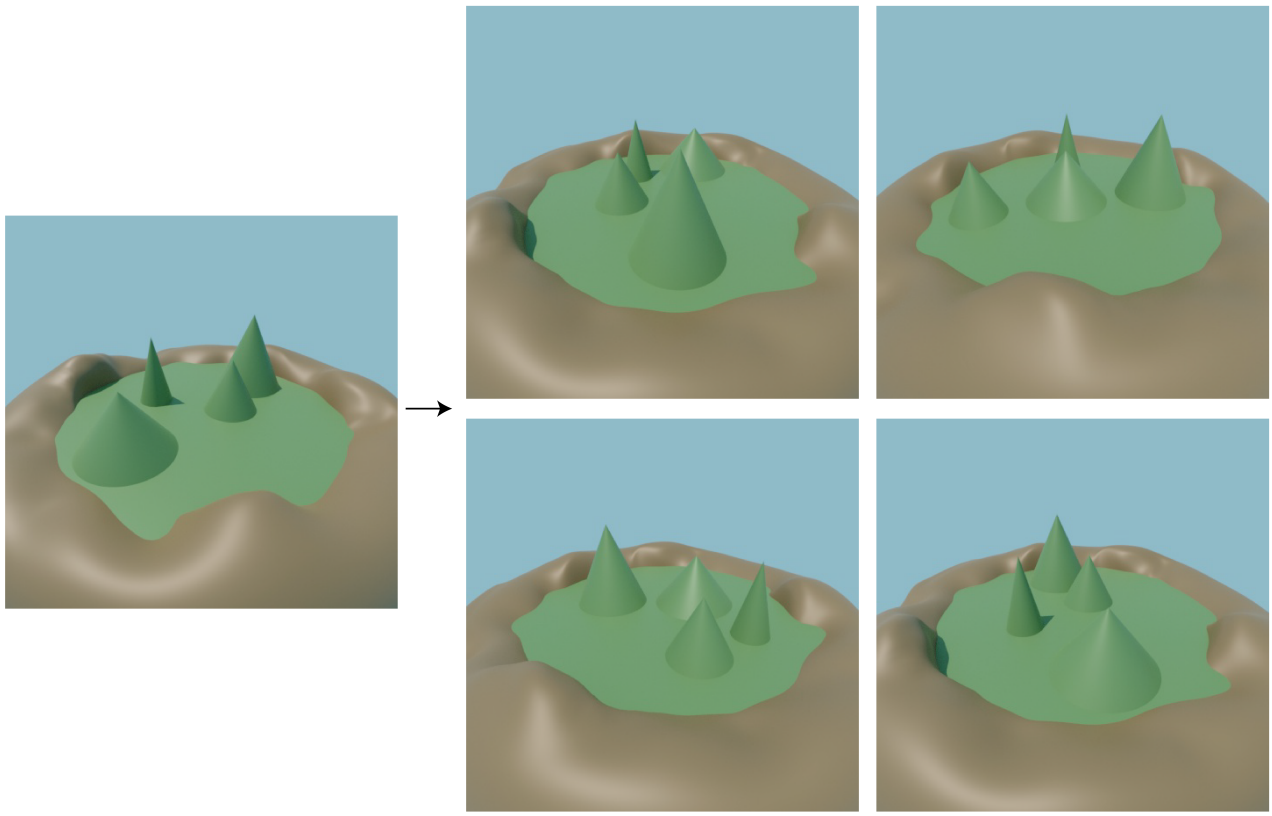

Figure S1. **Original task design, used to detect Alzheimer's disease**
In the 4-Mountains-Test [18], the participant first sees the image on the left and after a two-second delay is shown the images on the right, from which they need to choose the image which shows the same scene but from a different viewpoint. In all other images, the allocentric position of the mountains is changed (distractor images). The correct scene in this example is the lower right image.

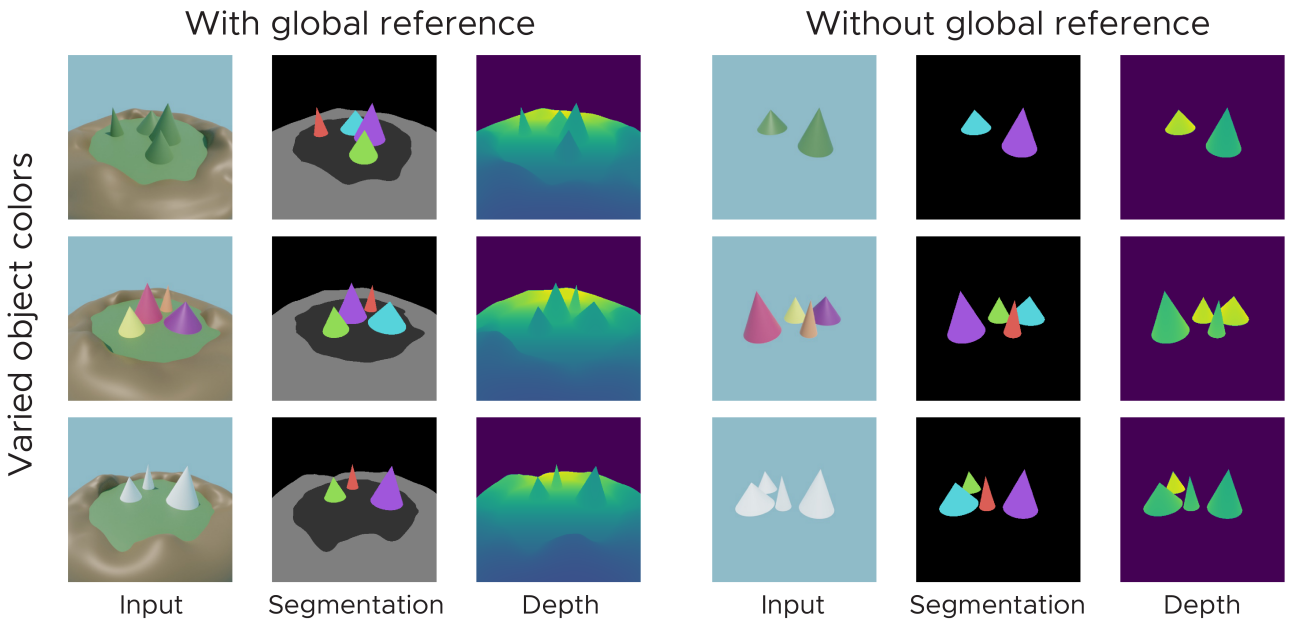

Figure S2. **Variations of allocentric scene perception (ASP) task**
We render six different variations of our dataset, by changing the colours of the object as well as the global reference. The colours are chosen to test model performance for segmenting objects with the same colour, objects which share the colour of the floor or are maximally different than the floor and objects which are already distinguishable by colour alone. We also vary the use of distal landmarks (global reference) to test model segmentation without additional cues regarding distances from boundaries in allocentric space.

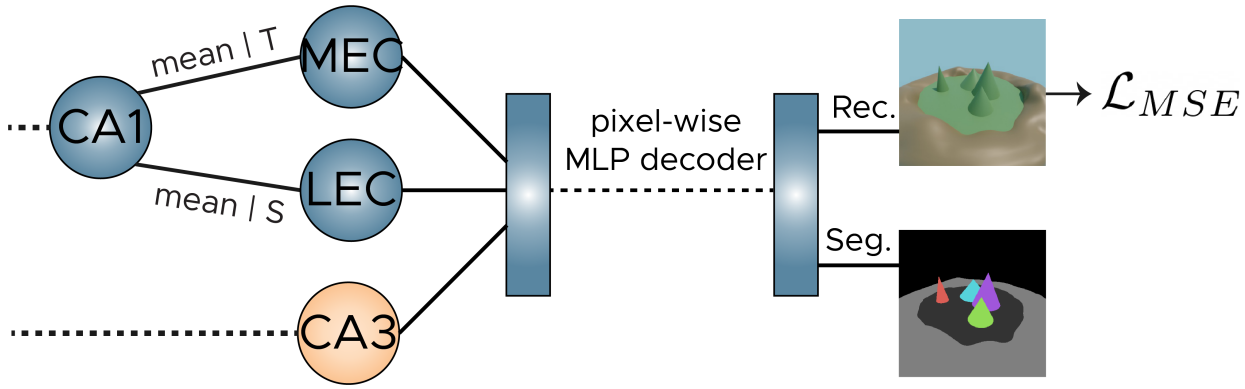

Figure S3. **Model architecture for reconstructing inputs**
When the model is tasked to reconstruct the input as well as segment objects, we use a pixel-wise decoder which takes input signals from MEC, LEC and CA3. MEC and LEC information is acquired by splitting the information into two pathways, one of which takes averaged time (T) information, thereby carrying information about space (MEC). In contrast, the other pathway takes averaged space (S) information, thereby retaining information about time (LEC). The MLP outputs four channels, where three are used for reconstructing the inputs, while the fourth is used for constructing segmentation masks.

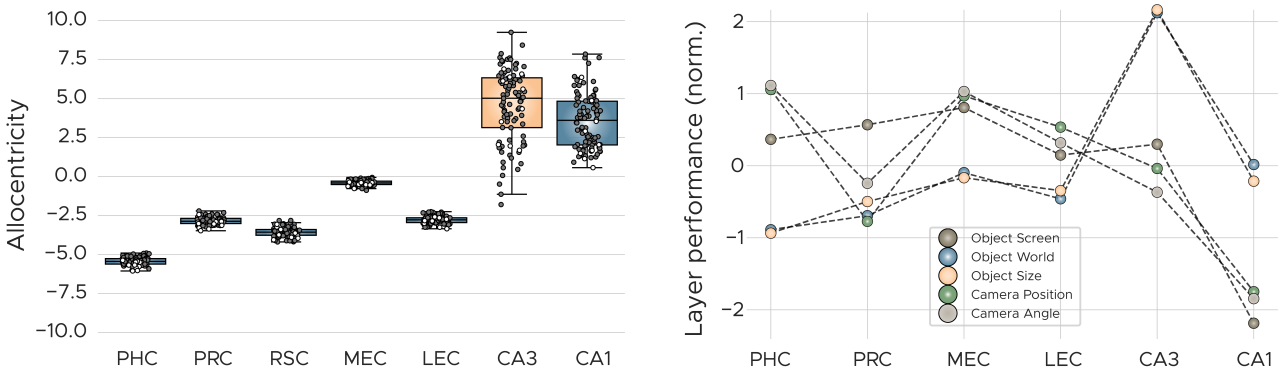

Figure S4. **Allocentricity score and layer readout**
(Left) We show the allocentricity score across layers, which measures the activity of neurons across different views within a scene, defined as the coefficient of variation of the activation of each artificial neuron across several images of the same scene. High values indicate that the neurons within this layer have similar activity profiles for images from the same scene independent of the viewpoint from where it was taken. Most world-centred responses are measured in late layers, indicating that hippocampal responses are similar for images from the same scene from different viewpoints. (Right) We use a linear readout across layers to investigate the information contained within them. Performance is normalized via z-scoring across layers.

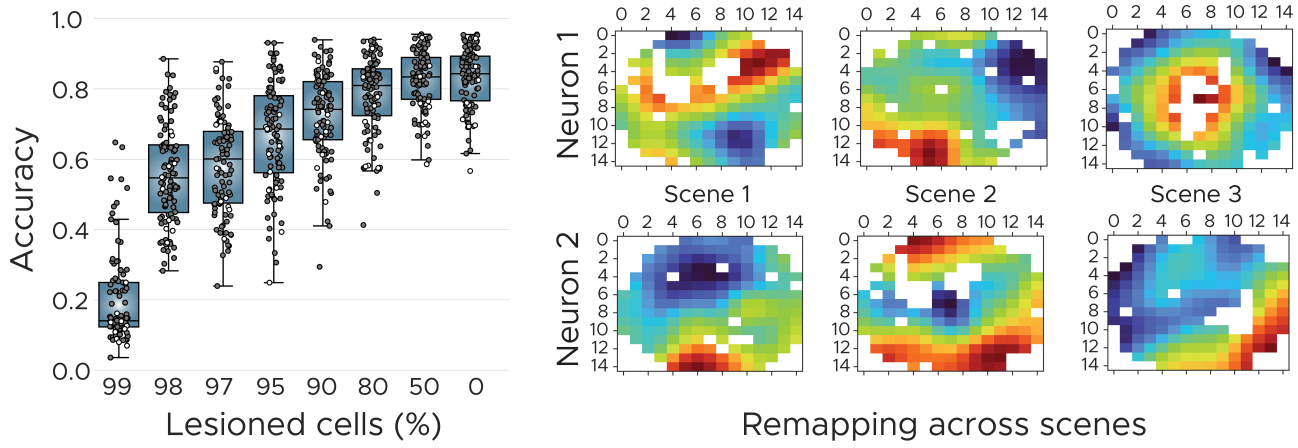

Figure S5. **Lesioning experiments and remapping across scenes**
(Left) Performance for differentiating scenes across the ratio of lesioned cells. (Right) Responses of two neurons across three different scenes, showing spatial selectivity remaps with scene identity.