

Processing of scene intrinsics in the ventral visual stream for object recognition

Shreya Kapoor

Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Bavaria, Germany

Bernhard Egger

Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany

Abstract

A hallmark of human vision is the ability to rapidly recognize objects in a complex naturalistic scene. However, the exact mechanisms behind the computational invariance of object recognition remain unknown. In this study, we investigate object constancy by estimating how the ventral visual stream processes shading, shadows, textures, and specularities. To accomplish this, we use object meshes from the Objaverse dataset to create distinct multiclass classification tasks. For every task, we render a dataset by excluding exactly one of the previously stated features at a time. Subsequently, we train a ResNet50 model on each dataset. The trained model is evaluated on Brain-Score; deviations in these metrics indicate the importance of a brain region in achieving invariance to a specific feature. A reduced score for a removed feature in a particular region implies its crucial role in processing that feature since the model classifies objects based on remaining scene intrinsics.