

Supplementary Information

Supplemental Movie

The whole-brain EEG-fMRI fusion movies can be viewed via <https://www.youtube.com/watch?v=8nc9BnuTosQ>. It provides a spatiotemporally resolved view of whole-brain activity at each millisecond. The color-coded values denote correspondences between fMRI and EEG responses in representational space (see Methods for details; 1,000 cluster-corrected sign permutation test, cluster-definition threshold $p < 0.001$, cluster threshold $p < 0.01$).

Supplemental Figures

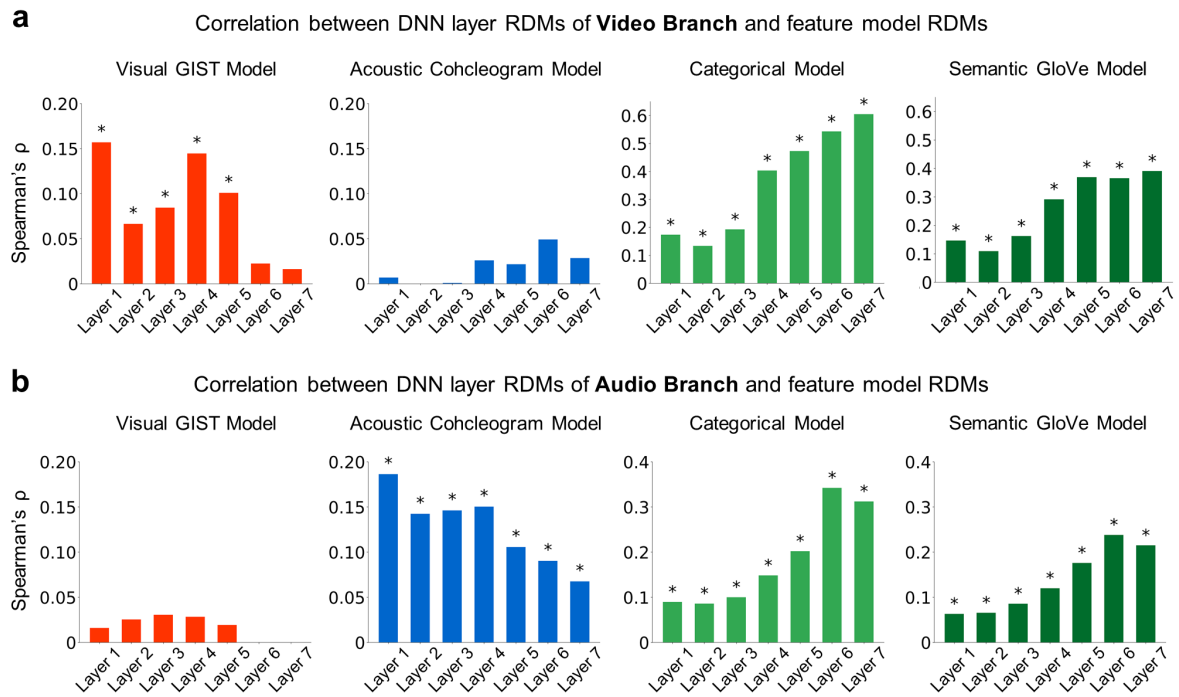


Figure S1: Correlation between layer RDMs of two DNN branches and feature model RDMs. The height of each bar denotes the Spearman's rank correlation value between two tested model RDMs. The asterisk above the bar represents significant correlations (10,000 permutation test with FDR correction at $p < 0.01$).

Supplemental Tables

Animals	Objects	Scenes
cat meowing	alarm ringing	beach
cow mooing	blender working	billiard hall
cricket chirping	book flipping	bowling hall
crown cawing	cellphone vibrating	church
dog barking	clock ticking	classroom
donkey braying	door opening	firework
duck quaking	drum rolling	lake
elephant trumpeting	glass clinking	music festival
frog croaking	heart monitor beeping	orchestra
goose honking	keyboard typing	raining
hen clucking	laptop opening	roller coaster
horse neighing	money counter working	soccer game stadium
lion growling	music box playing	stream
mouse squeaking	printer working	street with ambulance running
pig oinking	radio tuning	street with car horn
pigeon cooing	smoke detector beeping	street with car running
sheep bleating	telephone ringing	thunder and lightning
rooster crowing	toaster ding	traffic jam
sparrow cheeping	toilet flushing	train passing
wolf howling	zipper sliding	wind blowing

Table S1: Descriptions of video stimuli. Each line describes the content of one video stimulus.