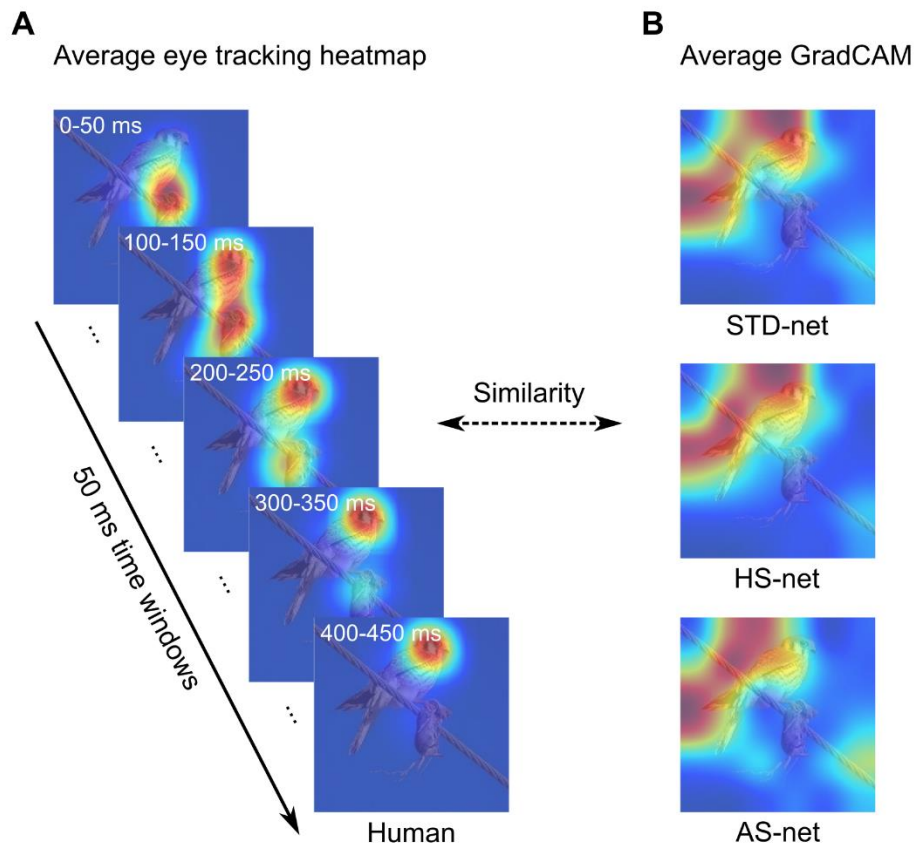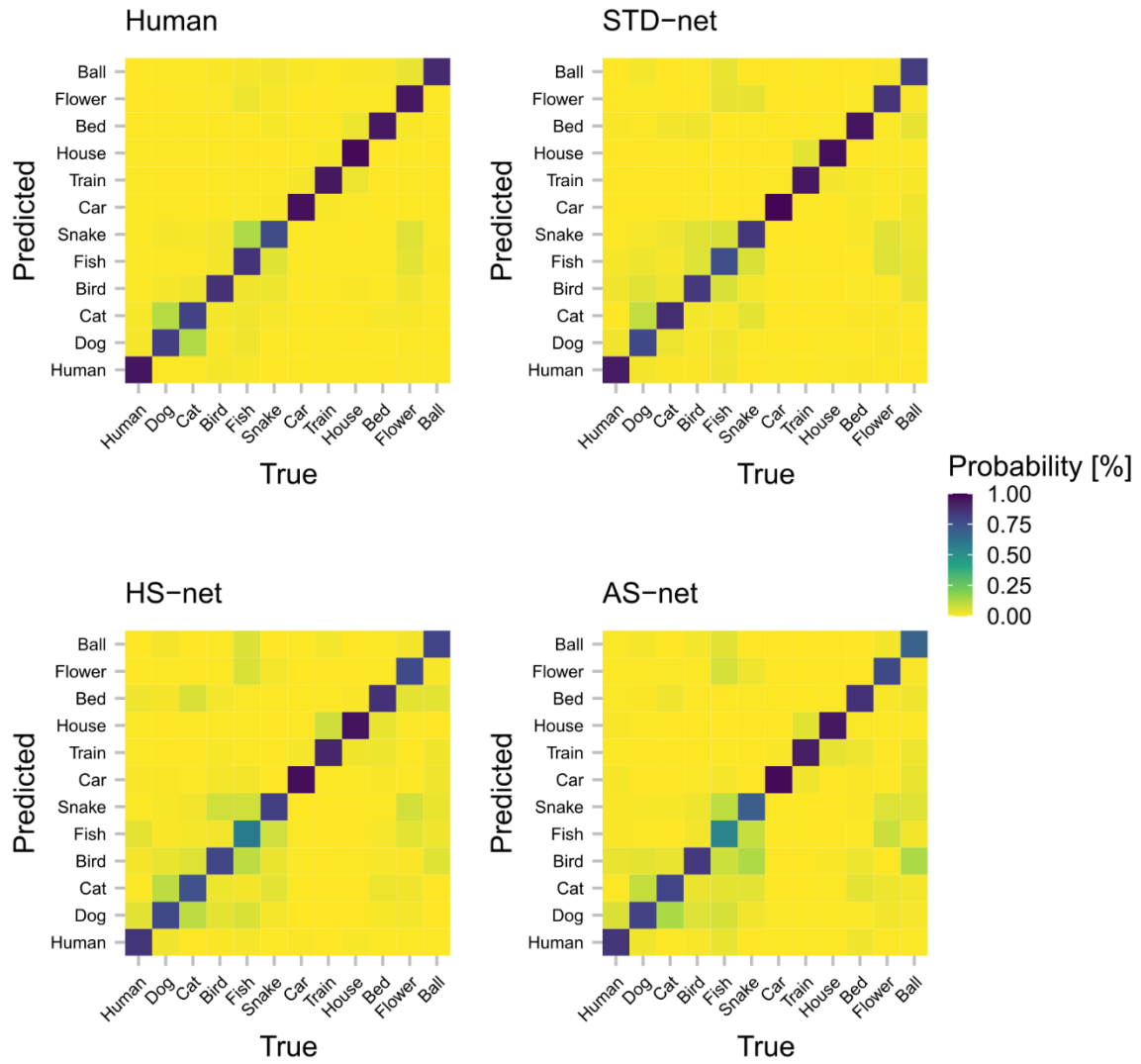# Supplementary Figures

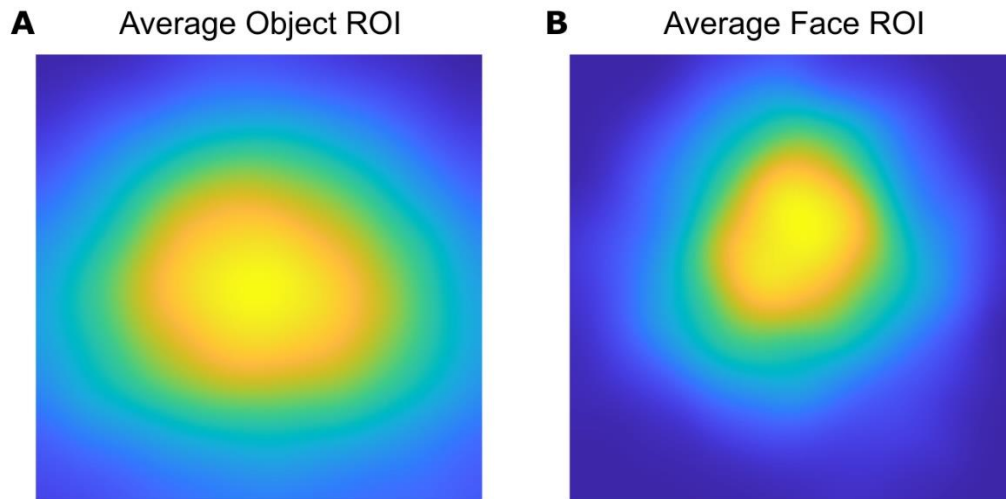**Supplementary Figure 1**
Image example with average visual attention of human participants and DCNN manipulation types. **(A)** Average human eye tracking heatmaps across 50 ms time windows. **(B)** Average DCNN GradCAM saliency maps for individual manipulation types.

Supplementary Figure 2

Confusion matrices displaying common miscategorizations by probability of occurrence for true versus predicted category combinations in human participants, standardly fine-tuned STD-nets, purposefully manipulated HS-nets, and inversely manipulated AS-nets across all testing images.

Supplementary Figure 3

Average manually segmented regions of interest (ROIs). **(A)** Average object ROI (N = 360). **(B)** Average face ROI (N = 133).