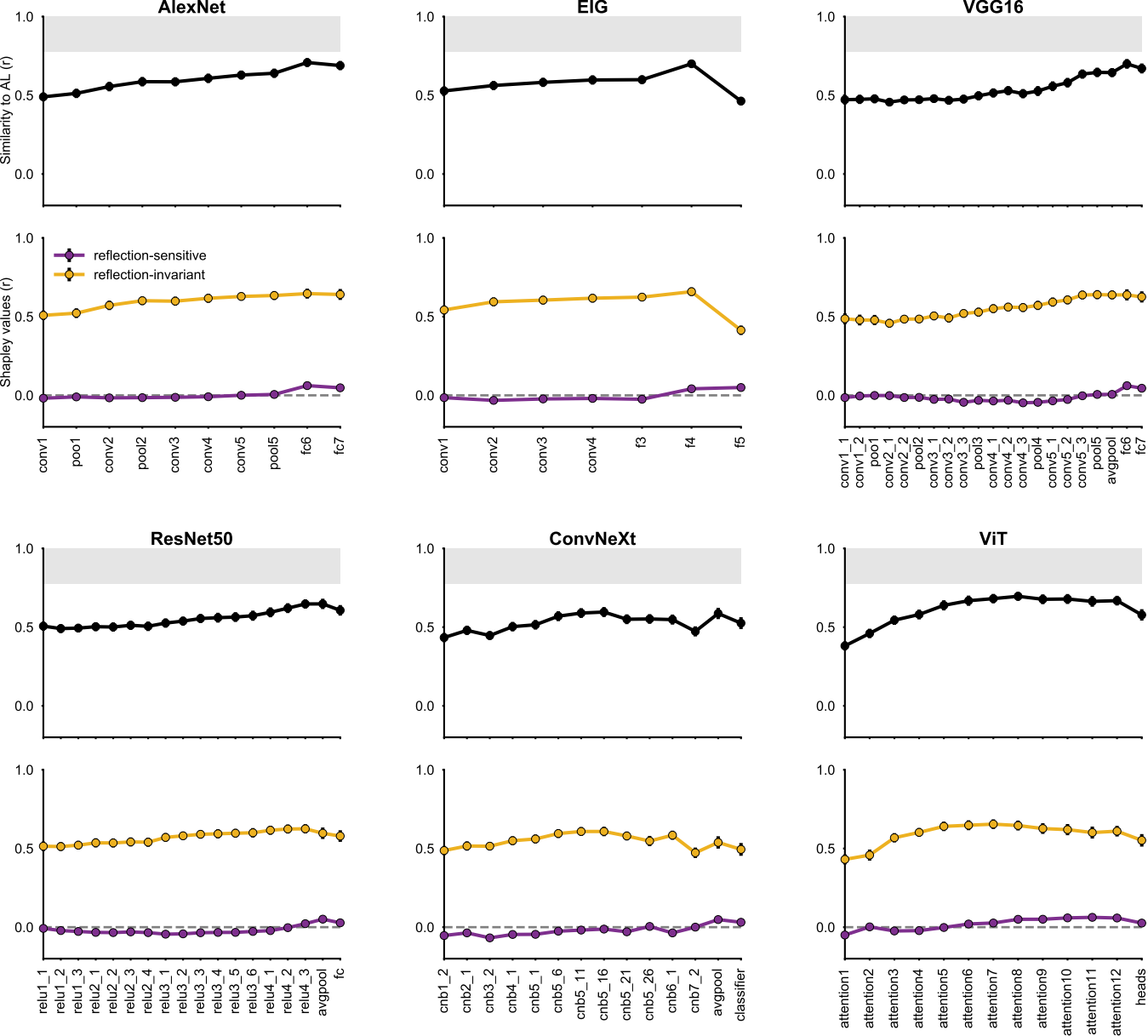
AlexNet
EIG
VGG16
ResNet50
ConvNeXt
ViT
Similarity to AL (r)
Shapley values (r)
reflection-sensitive
reflection-invariant
1.0
0.5
0.0
conv1
pool1
conv2
pool2
conv3
conv4
conv5
pool5
fc6
fc7
f3
f4
f5
conv1_1
conv1_2
pool1
conv2_1
conv2_2
pool2
conv3_1
conv3_2
conv3_3
pool3
conv4_1
conv4_2
conv4_3
pool4
conv5_1
conv5_2
conv5_3
pool5
avgpool
fc6
fc7
relu1_1
relu1_2
relu1_3
relu2_1
relu2_2
relu2_3
relu2_4
relu3_1
relu3_2
relu3_3
relu3_4
relu3_5
relu3_6
relu4_1
relu4_2
relu4_3
avgpool
fc
cnb1_2
cnb2_1
cnb3_2
cnb4_1
cnb5_1
cnb5_6
cnb5_11
cnb5_16
cnb5_21
cnb5_26
cnb6_1
cnb7_2
avgpool
classifier
attention1
attention2
attention3
attention4
attention5
attention6
attention7
attention8
attention9
attention10
attention11
attention12
heads