

Figure 1. (ImageNet Number of Groups vs. Error) SOP only needs 5 groups to achieve the same performance (error rate) as if it has 20 groups.

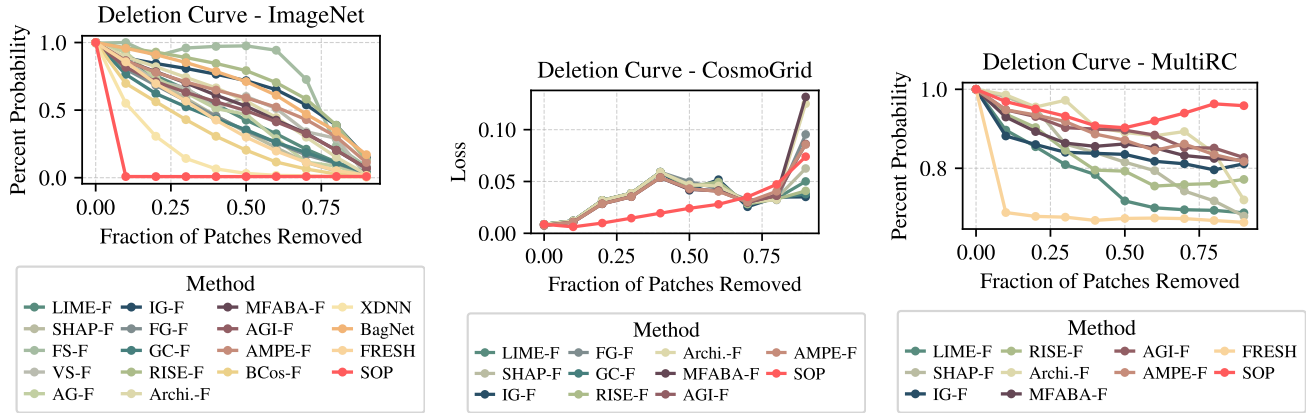


Figure 2. Deletion curves across datasets. SOP’s multi-group approach allows it to maintain low loss when only a few input features are preserved. Other methods show unstable behavior, especially in CosmoGrid.

Method	Error	IoU	Insertion	Deletion
FastSHAP	0.8270	0.0705	0.4556	0.7974
ViT-SHAP	0.7323	0.2447	0.5329	0.5715
AutoGnothi	0.7095	0.2229	0.5758	0.5337

Table 1. (ImageNet-S) Comparison of methods by error, IoU, insertion, and deletion metrics for Additional Baselines.

Computation Cost	Methods
1× forward pass	IG-F, GC-F, FG-F, BCos-F, XDNN, BagNet, FRESH
20× forward passes	LIME-F, SHAP-F, RISE-F, MFABA-F, AGI-F, AMPE-F, SOP
$\mathcal{O}(d^2)$ forward passes	Archipelago-F (due to pairwise interaction testing)

Table 2. Computation cost of different attribution methods in terms of number of forward passes.

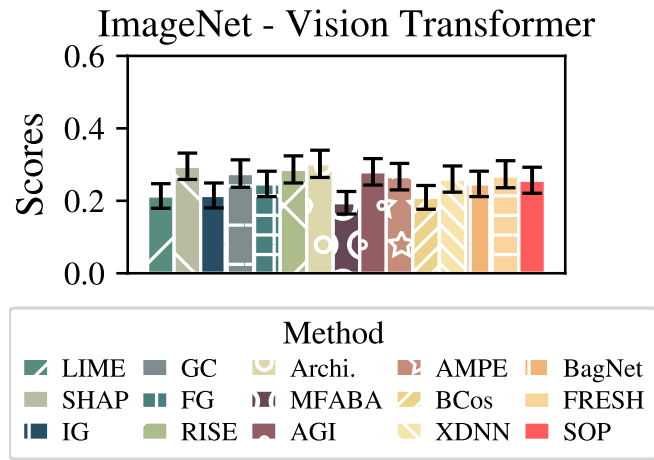


Figure 3. Human Distinction Task on ImageNet (50 examples). Most methods have similar human distinction task accuracy, including SOP.

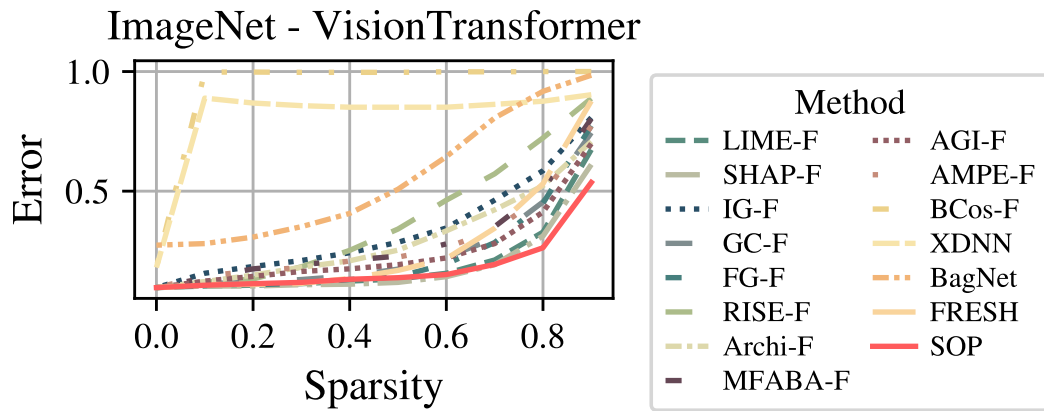


Figure 4. (ImageNet Sparsity) Larger figure for ImageNet Sparsity.