A   APPENDIX

A.1   IMAGENET PRE-TRAINING

In this section, we discuss the significance of no ImageNet pre-training used in our experiments. It has been common practice with pixel-wise SOD methods to initialize the feature extractor with pre-trained ImageNet weights (Liu et al., 2021a; Sandler et al., 2018; Lee et al., 2022). Since GNN based methods (Veličković et al., 2017; Dwivedi & Bresson, 2020), superpixel convolution (Zohourian et al., 2018b), and our proposed method SuperFormer do not have pre-trained ImageNet weights, we consider it unfair to compare the results between those with and without pre-trained weights since transfer learning is a huge advantage (Hendrycks et al., 2019; Zhuang et al., 2020) and not a major contribution to these works. Therefore, we take the reproducible codes provided by the authors of VST (Liu et al., 2021a), MobileNetV2 (Sandler et al., 2018), and TRACER (Lee et al., 2022) and train the model fully without initializing the feature extractor with ImageNet weights.