

## APPENDIX

## A. Visualization results on more datasets.

Here, Fig. A1 and Fig. A2 show the visualization results of the attention maps learnt by both the backbone method and the proposed method on the CUB and Stanford Cars datasets respectively.

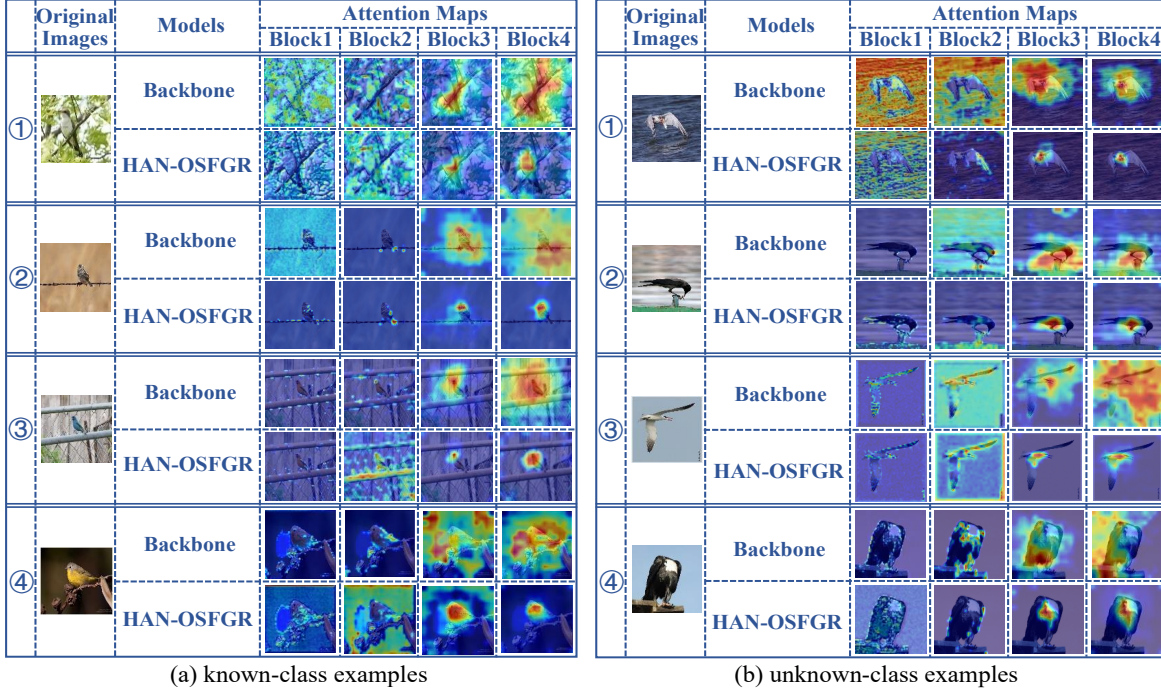


Fig. A1: Visualization of the attention maps learnt by both the backbone method (*i.e.*, the swin transformer) and the proposed method for (a) four known-class images and (b) four unknown-class images from the CUB dataset. All of these images are correctly predicted by HAN-OSFGR. The attentions in red regions are strongest, while the attentions in yellow, green, and blue regions decrease by degrees.

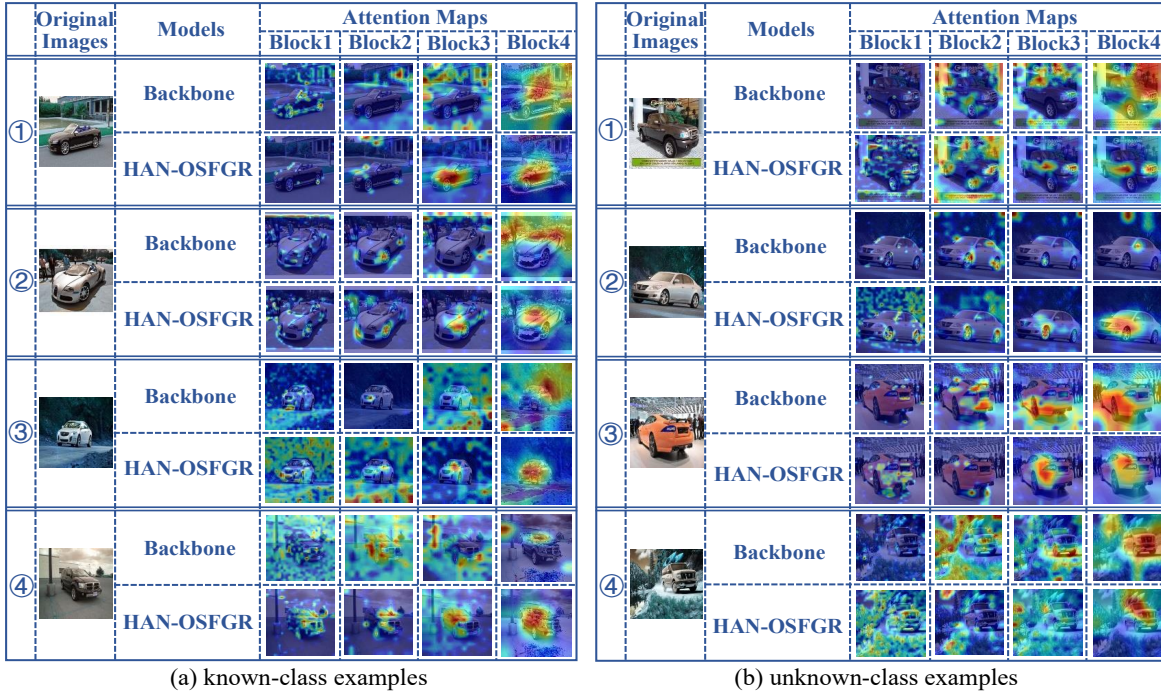


Fig. A2: Visualization of the attention maps learnt by both the backbone method (*i.e.*, the swin transformer) and the proposed method for (a) four known-class images and (b) four unknown-class images from the Stanford Cars dataset. All of these images are correctly predicted by HAN-OSFGR. The attentions in red regions are strongest, while the attentions in yellow, green, and blue regions decrease by degrees.