Label Relation Graphs Enhanced Hierarchical Residual Network for Hierarchical Multi-Granularity Classification

Jingzhou Chen[1], Peng Wang[1], Jian Liu[2], Yuntao Qian[1*]
[1]College of Computer Science, Zhejiang University, China
[2]Ant Financial Services Group
{11621038, pengwang18, ytqian}@zju.edu.cn, rex.lj@antgroup.com

1. Additional Experimental Results

1.1. Comparison with Conventional HMC Loss

To verify the effectiveness of the proposed hierarchical loss function, we compare it with the conventional HMC loss function that sums over binary cross-entropy losses from each hierarchical level in the same proposed network architecture. The conventional HMC loss ignores hierarchical relations between labels, and the post-processing step is needed to correct predictions that violate hierarchical constraints. Tab. 1 displays compared results on CUB-200-2011 with the relabeling proportion of 0%. The results with our loss are about 8% and 4% higher than those with the conventional HMC loss in OA and $AU(\overline{PRC})$, respectively. These results verify the effectiveness of the proposed hierarchical loss that formalizes the semantic relations between labels in the hierarchy into a tree.

Table 1. OA on the species level and $AU(\overline{PRC})$ results with the relabeling proportion of 0% by comparing our combinatorial loss with the conventional HMC loss.

Loss	Conventional HMC Loss	Our Combinatorial Loss
OA	77.90	**86.60**
$AU(\overline{PRC})$	0.921	**0.969**

1.2. Visual Demonstration

We conduct visualization experiments to demonstrate that granularity-specific blocks can capture different regions of interest while hierarchical knowledge can be transferred across levels via hierarchical residual interaction and supervision of combinatorial loss. To this end, we adopt Grad-Cam to visualize different attention regions of each hierarchical level by propagating their respective gradients back to feature maps generated from the trunk network. Fig. 1 illustrates two species of birds: House Wren and Marsh Wren from the same family (Troglodytidae) and order (Passeriformes). Domain knowledge reveals that Marsh Wrens have a more distinctive eyebrow than House Wrens. In Fig. 1, the order level focuses on the head and legs, and the family level pays close attention to the head and feathers on the wing and tail. In addition to feathers, the species level concentrates on the head. Three hierarchical levels give similar attention to the head with a distinctive eyebrow, while they have unique regions of interest.

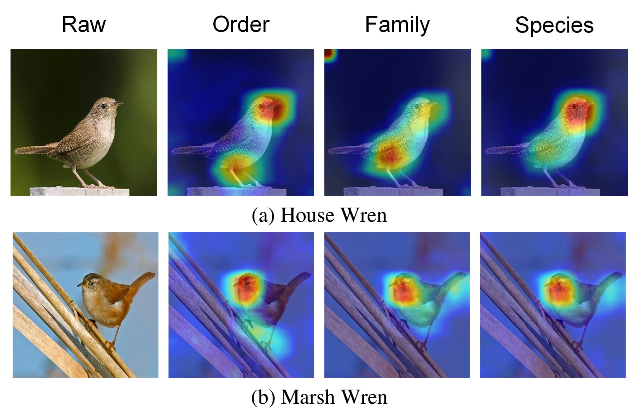
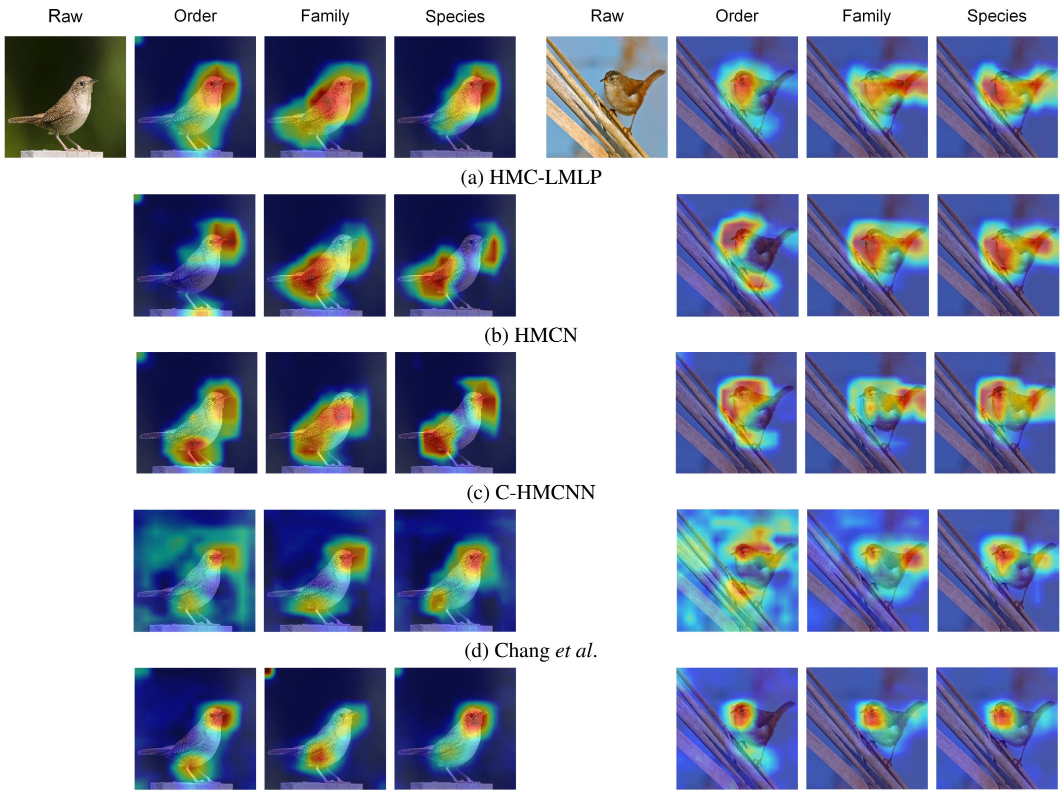


Figure 1. Visual attention maps of two species of birds from the same family (Troglodytidae) and order (Passeriformes).

1.3. Visual Comparison with State-of-the-art Methods

We compare attention regions of our proposed method to state-of-the-art methods: HMC-LMLP, HMCN, C-HMCNN, and the method proposed by Chang *et al*. to analyze regions of interest corresponding to different hierarchical levels using Grad-Cam. Each hierarchical level propagates their respective gradients back to feature maps produced from the trunk network (ResNet-50). We visualize results of relabeling proportion 0% without reducing the image resolution and illustrate two bird species

*Corresponding author



(a) HMC-LMLP

(b) HMCN

(c) C-HMCNN

(d) Chang *et al*.

(e) Ours

Figure 2. Visual attention maps of compared state-of-the-art methods and ours from three hierarchical levels on House Wren (left) and Marsh Wren (right) belonging to the same family (Troglodytidae) and order (Passeriformes).

House Wren and Marsh Wren belonging to the same family (Troglodytidae) and order (Passeriformes) from CUB-200-2011 in Fig. 2. According to the bird's domain knowledge, Marsh Wrens have paler underparts than House Wrens with a more distinctive eyebrow. As shown in Fig. 2, all approaches pay attention to similar distinctive regions like the head, legs, and features on the wing and tail, while our approach is more focused on the target across all three hierarchical levels.