

王舜. 基于物种分类树的野生动物监测图像层次化分类方法[J]. 智能计算机与应用, 2025, 15(1): 81-87. DOI: 10.20169/j.issn.2095-2163.24061102

基于物种分类树的野生动物监测图像层次化分类方法

王舜

(北京林业大学 工学院, 北京 100083)

摘要: 针对野生动物监测图像物种分类模型结果复核成本高昂的问题, 本文提出一种基于物种分类树的层次化分类方法。该方法在物种分类树的纲、目、科、属和种 5 个分类层级上进行零样本分类, 通过提供更丰富的物种判定信息, 降低复核模型结果的人工成本; 利用类别间的层次关系, 引入软决策和路径矫正策略, 在细粒度类别上分类错误时, 模型提供的粗粒度结果也具有参考价值。与路径矫正策略增强的基线方法相比, 该方法在 5 个分类层级上的准确率分别提升了 0.27%、1.44%、1.42%、1.39% 和 1.03%, 并将错误严重程度降低了 1.1%。本文提出的方法提高了层次化分类的准确率和一致性, 增强了模型输出结果的可解释性, 降低了人工复核成本, 有利于提高生态学家对深度学习模型的信赖程度和使用意愿。

关键词: 野生动物监测; 物种分类; 零样本分类; 层次化分类

中图分类号: TP391

文献标志码: A

文章编号: 2095-2163(2025)01-0081-07

Hierarchical classification method for wildlife monitoring images based on species classification tree

WANG Shun

(School of Technology, Beijing Forestry University, Beijing 100083, China)

Abstract: To address the problem of high cost in the review of model results in species classification in wildlife monitoring images, this work proposes a hierarchical classification method based on the species classification tree. This method simultaneously performs zero-shot classification at the five classification levels of Class, Order, Family, Genus and Species in the species classification tree and reduces the labor cost of reviewing model results by providing more information about species identification. The proposed method utilizes hierarchical relationships between categories and introduces soft decision and path correction strategies. The coarse-grained results provided by the model are also valuable, while classification fails at the fine-grained species level. Compared to the baseline method enhanced by the path correction strategy, the accuracy of the proposed method increased by 0.27%, 1.44%, 1.42%, 1.39%, and 1.03%, respectively, at each classification level, and the mistake severity was reduced by 1.1%. The proposed method improves the accuracy and consistency of hierarchical classification, enhances the interpretability of classification results, reduces the labor cost of reviewing model results, and increases ecologists' confidence and willingness to use deep learning models.

Key words: wildlife monitoring; species classification; zero-shot classification; hierarchical classification

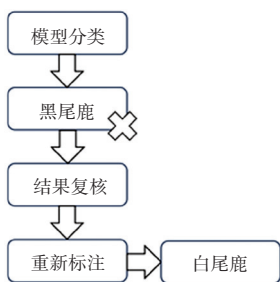
0 引言

深度学习在计算机视觉领域的巨大成功, 激发了研究人员将其应用于野生动物监测图像的兴趣^[1]。近年来, 机器学习尤其是深度学习在监测图像的各种计算机视觉任务上取得了良好的进展, Chen 等^[2]首次利用深度神经网络实现野生动物监测图像物种分类; 宫一男等^[3]实现了东北虎图像的自动识别; Chu 等^[4]结合地理位置信息, 提高了野生

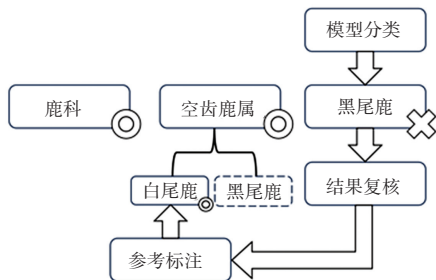
动物监测图像物种分类性能; Norouzzadeh 等^[5]在分类的基础上实现了野生动物的个体计数; Yu 等^[6]进行了动物姿态估计方面的研究。这些方法为理解特定监测区域的物种分布、种群规模和动物行为提供了数据支撑, 为生态研究中物种多样性、相对丰度和环境承载力等关键生态指标的评估奠定了基础。

但在生态研究中, 需要对野生动物监测图像的分类结果进行复核, 以保证生态数据的准确性, 而错误分类结果的复核过程无异于重新标注, 依赖于专

业知识并且耗时低效。通常模型只在一个分类层级(如物种层级)进行分类,如果模型给出错误结果,标注人员就需要重新确认样本的所属类别,如图1(a)所示;而层次化分类则同时给出多个层级的结果,在分类错误时也能提供更多有利于确认物种的信息,从而降低对标注人员领域知识的依赖,提高模型的可用性。如图1(b)所示,当模型在物种层级给出错误的预测结果黑尾鹿时,标记人员可以参考科级、属级的结果,快速地从相近物种中确认正确类别为白尾鹿。此外,将人工智能领域的算法和模型应用到生态领域,需要获取生态社区的信任^[7]。层次化分类有利于降低模型的错误严重程度,减少不合理的输出结果(例如把陆生动物判定为水生动物),从而维护用户对模型的信心^[8]。同时,层次化分类给出更明确的分类路径,利用领域知识解决领域问题,为模型输出提供了更好的可解释性,使生态研究人员更容易接纳深度学习模型。



(a) 单层级分类结果复核过程



(b) 层次化分类结果复核过程

图1 两种分类方式的结果复核过程

Fig. 1 Result review of the two classification methods

Silla 等^[9]的综述表明,现有研究集中于将监督学习模型从单层级分类扩展到多层级分类。而对于开放词汇模型如 CLIP (Contrastive Language-Image Pre-training) 的层次化分类研究较少^[10]。本文将开放词汇模型显式地扩展到层次化分类,讨论了零样本设置中野生动物监测图像的层次化物种分类。

1 相关工作

类别概念之间存在天然的层次关系,例如“动物、鸟类、鸽子”就是一组具有从属关系的、粒度由粗到细的类别。层次化图像分类 (Hierarchical Image Classification) 旨在使模型学习层次关系,对图像在多个粒度的标签层级上进行分类,相比于单层级分类,层次化分类能够提升性能^[9,11]。目前层次化图像分类的研究可分为两方面:学习标签层级的具体优化目标和层次化模型结构设计。在优化目标方面,Bertinetto 等^[8]提出了层次化的交叉熵损失和反映标签间语义关系的标签编码;Wu 等^[12]也通过设计语义感知的损失函数降低了分类错误的严重程度;Wang 等^[13]利用标签层级设计损失矩阵,学习类别之间的层次关系。在模型结构方面,Yan 等^[14]提出利用分支结构增强层次化分类模型;Inoue 等^[15]及 Zhang 等^[16]提出连接不同分支的特征,从而联系各个层级的分类任务;Chang 等^[17]解耦多层级的图像特征,发现解耦后不同粒度的任务可以相互促进。对于开放词汇分类模型,CHiLS^[18]利用大语言模型为粗粒度标签生成细粒度标签,将粗粒度分类任务由细粒度任务代理,从而提升分类效果;Ren 等^[19]也探究了利用大语言模型为开放词汇模型生成文本提示的层次化分类方法。

2 本文方法

2.1 零样本层次化物种分类

为了保证类内的视觉相似性以及更具实际意义的层级上进行分类,本文参照物种分类学,选择纲 (Class)、目 (Order)、科 (Family)、属 (Genus) 和种 (Species) 5 个分类层级进行研究。

1) 根据类别间从属关系设计文本提示

由于层次化分类可以视为多个单层级分类任务的组合,所以提高单层级分类性能的方法也可以纳入层次化分类方法中。本文采用提示工程的方法,通过引入相邻分类层级的文本描述提高图像与各层级的文本的匹配程度。本文的文本提示模板见表1,其中种级的“{class}”由英文物种名替换,其余层级则使用拉丁文学名,“{subclasses}”表示下层类别的物种名,“{superclass}”表示相邻的上一层级的物种名。相邻层级类别名称的加入,建立了不同层级类别之间的联系,从而使模型能依据类别间的层次结构进行分类。

表 1 引入层次结构的文本提示模板

Table 1 Text prompt template introducing hierarchical structure

分类层级	文本提示模板
纲	a photo of a {class}, such as {subclasses}
目、科、属	a photo of a {class}, a kind of {superclass}, such as {subclasses}
种	a camera trap image of a {class}, a kind of animal

2) 基于物种分类树的自上而下的层次化分类

生物分类学按照树状层次结构组织物种类别,每个物种在不同层级上都有相应的分类。沿着分类树从根结点到叶结点的路径,物种的分类逐渐细化,类别粒度由粗到细,语义从抽象到具体。更高层级的特征因其一般性更容易识别,因此可以根据分类树的结构,采取由粗到细的逐层分类策略,利用粗粒度层级的分类结果排除细粒度类别中的错误选项,从而提高物种分类的准确性。本文提出了基于物种分类树的自上而下的层次化分类方法,用于野生动物监测图像的物种分类。该方法从更一般的类别逐渐细化到更具体的类别,生成包含多层次信息的、更具可解释性的分类结果。自上而下的层次化分类具体实现过程由算法 1 描述,与 CLIP 模型相同,图像特征与文本特征的相似度分数由余弦相似度计算。

算法 1 基于物种分类树的自上而下层次化分类方法

输入 isa_mask, 维度为 $N \times N$ 的二值矩阵。记录 N 个类别间折从属关系, isa_mask[i][j] 为真表示第 j 个类属于第 i 个类的子类, layer_mask, 维度为 $L \times N$ 的二值矩阵。记录 N 个类别 L 个层级的从属关系, layer_mask[i][j] 为真表示第 j 个类是第 i 层的类别。img 表示输入图像, text 表示所有 N 个类别文本提示。

输出 labels, 长度 L 的数组, 表示自上而下的各层级的分类结果。

1. $I, T = \text{Encoder}_I(\text{img}), \text{Encoder}_T(\text{text})$ ▷ 特征提取

2. $\text{logits} = \text{similarity}(I, T)$ ▷ 计算相似度分数

3. $\text{pred} \leftarrow \text{null}, \text{labels} \leftarrow \emptyset$ ▷ 变量定义

4. **for** $i \leftarrow 1$ **to** L **do**

5. $\text{measked_logits} = \text{logits} * \text{layer_mask}[i]$ ▷ 取第 i 层类别相似度分数

6. **if** $i > 1$ **then**

7. ▷ 取得上一层级分类结果所有子类的相似度分数, 其余类别置 0

8. $\text{masked_logits} = \text{logits} * \text{isa_mask}[\text{pred}]$

9. **end**

10. $\text{pred} = \text{argmax}(\text{masked_logits})$ ▷ 得到第 i 层的分类结果

11. $\text{labels} = \text{labels} \cup \text{pred}$ ▷ 将结果加入返回值

12. **end**

13. **return** labels ▷ 返回 L 个层级上的分类结果

野生动物监测图像的层次化物种分类不仅能够提高模型的准确性, 还能够在结果中提供更多有助于理解和验证的上下文信息, 有助于减少因细粒度分类错误导致的潜在问题, 同时也提供了更为丰富和透明的决策支持, 从而提高模型的可用性。

2.2 分类路径的纠正策略

与剪枝类似, 自上而下的层次化分类在分类树中预先排除更深层级的错误选项, 逐步缩小标签空间。然而排除的类别是否合理取决于当前分类结果的正确性, 因此早期的分类错误会累积到更深的层级, 导致下一层级的准确率不可能高于上一层级。随着层级加深, 分类错误的指数级累积严重限制了深层级分类的准确率。例如, 即使每个分类层级相对于上一层级的准确率都高达 90%, 进行 5 个层级的分类后的准确率(即 0.9^5)也低于 60%。为使自上而下的层次化分类更加实用, 需要引入纠正前期的分类错误的策略。

由以上分析, 算法 1 的缺陷在于过早地做出决策, 基于当前结果完全地排除了一些类别在更深层级中进行匹配的可能。因此本文提出软决策, 利用上一层级的概率分数对当前层级 logits 分数进行修正, 保留但削弱先前层级的决策结果的影响, 避免直接缩小分类的范围, 使更深层级的决策有纠正早期错误的机会。使用软决策的分类概率由下式计算:

$$\text{score}_l = \text{Softmax}(\text{map}(\text{score}_{l-1}, l) \cdot \text{logits}_l / \tau) \quad (1)$$

其中, l 表示当前分类层级, $\text{score}_0 = \mathbf{1}_N$, logits_l 表示经掩码运算的第 l 层类别相似度分数的向量, 维度为 $1 \times N$, $\text{map}(\cdot)$ 用于将上一层级的分数 score_{l-1} 映射到第 l 层的子类别的位置上, 从而与该层的 logits 分数对应相乘, 温度 τ 在此处称为独立系数, 用于控制上一层级的分类概率对当前层级分类

的影响程度。

当独立系数极大时, score 为均匀分布, 各个层级独立决策, 上层分类结果不对后续分类产生影响; 而独立系数非常小时, 每层的 score 接近于 one-hot 向量, 将直接改变下一层级的分类范围, 软决策退化为自上而下的层次化分类。假设每个层级对后续层级的分类应保持相同的影响程度, 本文在各层级的软决策中使用相同的独立系数, 同时对各层级的 logits 进行归一化, 以消除数值量级上的差异。

软决策为层次化分类提供了在更深层级纠正先前层级的决策错误的可能性, 从而提高准确率。然而软决策不修正先前层级的错误结果, 这使模型的输出失去层次关系的一致性, 为解决这一问题, 需要对分类路径进行适当的矫正。假设在软决策过程中, 模型通过综合考虑各层级的结果, 能够在最终层级做出更准确的预测, 那么将分类路径自下而上地进行修正, 可以在层次一致性和分类准确率之间实现较好的平衡, 从而提高模型的输出质量。

基于软决策和路径矫正的层次化物种分类过程如算法 2 所示, PathCorrection 函数实现了路径矫正过程, 即根据软决策的最终结果映射到各个层级, 得到最终的层次化分类结果。

算法 2 基于软决策和路径矫正策略的层次化分类算法

输入 isa_mask, 维度为 $N \times N$ 的二值矩阵。记录 N 个类别间折从属关系, isa_mask[i][j] 为真表示第 j 个类属于第 i 个类的子类, layer_mask, 维度为 $L \times N$ 的二值矩阵。记录 N 个类别与 L 个层级的从属关系, layer_mask[i][j] 为真表示第 j 个类是第 i 层的类别。img 表示输入图像, text 表示所有 N 个类别文本提示, τ 表示软决策中的独立系数。

输出 labels, 长度为 L 的数组, 表示自上而下的各层级的分类结果。

1. $I, T = \text{Encoder}_1(\text{img}), \text{Encoder}_2(\text{text})$ ▷ 特征提取

2. $\text{logits} = \text{similarity}(I, T)$ ▷ 计算相似度分数

3. $\text{score} \leftarrow \text{null}, \text{labels} \leftarrow \emptyset$ ▷ 变量定义

4. **for** $i \leftarrow 1$ **to** L **do**

5. $\text{measked_logits} = \text{logits} * \text{layer_mask}[i]$ ▷ 取第 i 层的类别

6. **if** $i > 1$ **then**

7. ▷ 根据上一层级的概率修正当前层级的 logits 分数

8. ▷ map 将概率分数映射到第

i 层的类别上

9. $\text{masked_logits} = \text{masked_logits} * \text{map}(\text{score}, i)$

10. **end**

11. $\text{score} = \text{Softmax}(\text{masked_logits} / \tau)$ ▷ 由独立系数 τ 计算概率分数

12. $\text{labels} = \text{labels} \cup \text{argmax}(\text{score})$ ▷ 得到第 i 层的分类结果

13. **end**

14. $\text{labels} = \text{PathCorrection}(\text{labels})$ ▷ 由最终结果恢复分类路径

15. **return** labels ▷ 返回 L 个层级上的分类结果

3 实验与讨论

3.1 实验设置及评价指标

1) 层次化分类数据集

Koh 等^[20]提出了重要的公开野生动物监测数据集 iWildCam。本文在此基础上抽取了一个包含 36 个物种的样本数量均衡的子集 iWildCam36, 针对纲、目、科、属和种 5 个层级进行层次化分类研究, 并从维基百科获取每个物种在上述分类层级的标签, 将单层级的 iWildCam36 数据集转换为层次化分类数据集。iWildCam36 包含哺乳动物纲 (Mammalia) 和鸟纲 (Aves) 及下属的 8 个目、21 个科、34 个属的 36 种野生动物, 每类约 100 张图像。类别标签的层次结构如图 2 所示, 由于排版限制, 除种级标签外, 其余层级标签均在图中标出。

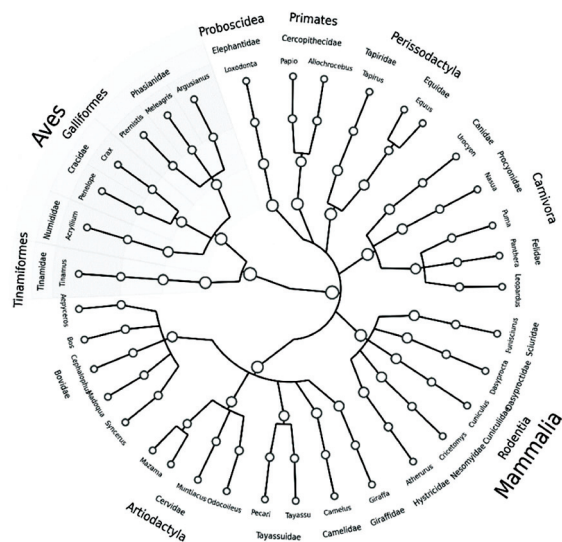


图 2 iWildCam36 数据集标签的层次结构

Fig. 2 Hierarchy of iWildCam36 dataset labels

2) 评价指标

本文通过层级准确率、层次一致性和错误严重程度 (Mistake Severity) 对提出的层次化分类方法进行

行评估,并与零样本基线进行比较。层级准确率即各个分类层级单独计算的准确率。层次一致性评估多层级结果中在分类树中构成分类路径的比例。错误严重程度即模型输出结果到真实结果之间的距离,与监督学习模型的层次化分类方法类似^[8]。本文利用预测类别和标签类别在分类树中的最近公共祖先(Lowest Common Ancestor, LCA)结点的高度度量错误严重程度,计算过程如下式:

$$\text{mistake severity} = \text{mean}(\text{height}(LCA(p, l)) / H) \times 100\% \quad (2)$$

其中, p 表示模型在分类树的最深层级的分类结果; l 表示标签真值; H 表示分类树的高度; $\text{height}(\cdot)$ 表示求取类别结点在树中高度。

3.2 结果及分析

基于 CLIP 模型,本文提出的层次化分类方法

与在多个分类层级上分别进行开放词汇零样本分类的基线方法的性能对比见表 2,其中独立系数 τ 取值为 0.04。相比于基线方法,本文方法提升了各个层级尤其是目、科和属层级的分类准确率,并在自上而下的逐层分类中将种级的准确率提高了 1.03%,即使与路径矫正策略增强的基线相比,本文方法仍然在各层级准确率上带来提升,这说明本文所提出的方法受益于类别间的层次关系并有效缓解了错误的逐层积累;另一方面,尽管基线方法能在多个层级上分别进行分类得到层次化的结果,但无法保证模型在相互独立的各层级分类中保持一致,同时科级和属级较低的准确率进一步限制了层次一致性。与之相反,本文提出路径矫正策略使模型输出始终符合分类树逻辑,还将模型分类错误的严重程度降低 1.1%,使模型的输出更接近正确结果。

表 2 所提出的方法与基线方法的性能对比

Table 2 Performance comparison between the proposed method and baselines

方法	各物种分类层级准确率					层次一致性	错误严重程度
	纲	目	科	属	种		
多层级分类(基线方法)	95.89	48.31	40.28	24.39	56.36	23.78	62.42
+路径矫正(增强基线)	95.67	81.56	70.97	59.25	56.36	100.00	62.42
层次化分类(本文方法)	95.94	83.00	72.39	60.64	57.39	100.00	61.32

3.3 消融实验

1) 本文方法各组成部分的消融

首先对本文方法的各个组成部分进行消融,以验证各组件的有效性,消融结果见表 3。层次文本提示改善了所有评价指标;基于物种分类树的自上而下分类保证了分类结果的一致性,但显著降低了种级准确率并加剧了错误严重程度,这表明在早期的分类中出现了更多错误;其次,软决策以层次一致性为代价,有效地利用了层次化分类提高种级的准

准确率;最后,路径矫正策略降低了错误严重程度并恢复了层次一致性,相对于软决策,目、科和属级的结果分别提高了 14.75%、30.47% 和 28.53%,这与现有研究的结论相符,即为粗粒度类别构造细粒度的分类层级,然后将细粒度分类结果映射回粗粒度类别能够提升粗粒度分类性能^[18]。本文方法在此基础上合理利用了粗粒度的分类结果参与细粒度分类,进一步提升了细粒度的种级分类任务性能。

表 3 所提出方法各组成部分的消融实验

Table 3 Ablation study on each component of the proposed method

方法	各物种分类层级准确率					层次一致性	错误严重程度
	纲	目	科	属	种		
多层级分类基线	95.89	48.31	40.31	24.39	56.36	23.78	62.42
+层次文本提示	96.81	68.25	41.56	31.75	56.86	31.11	61.82
+自上而下分类	96.81	68.67	39.53	31.50	29.36	100.00	66.29
+软决策	96.81	68.25	41.92	32.11	57.39	33.06	61.32
+路径矫正	95.94	83.00	72.39	60.64	57.39	100.00	61.32

2) 分类层级与独立系数的消融

减少层级数量同样可以缓解分类错误的逐层累

积,本文将分类层级减少至 3 层(纲、目和种)和 2 层(目和种)以探究层级数量对性能的影响,同时调

整独立系数适应层级数量的变化,结果见表4。由表4可知,增大独立系数会降低在全部5个层级进行分类的准确率,但能提高仅使用3个或2个层级时的分类效果。增大独立系数会降低上一层级分类结果对下一层级的影响程度,因此在准确率较低的层级参与层次化分类时,可能需要更低的独立系数

以保证模型在这些层级上的分类性能,以免逐层累积的错误对最终层级的影响超过类别层次结构带来的增益。在仅使用纲、目和种3个层级进行层次化分类, τ 取值为0.4时,种级准确率相比基线提高了2.7%,错误严重程度下降2.87%。

表4 分类层级与独立系数的消融

Table 4 Ablation of classification levels and independent coefficient

实验设置	各物种分类层级准确率					错误严重程度
	纲	目	科	属	种	
多层级分类基线	95.89	48.31	40.31	24.39	56.36	62.42
$\tau = 0.04$	95.94	83.00	72.39	60.64	57.39	61.32
	96.25	83.39			57.53	60.97
$\tau = 0.08$		83.39			57.50	60.98
	96.17	83.00	72.22	60.33	57.08	61.14
	96.17	83.08			57.28	61.12
$\tau = 0.1$		84.14			58.11	60.42
	96.33	83.17	72.25	60.53	57.25	61.04
	96.89	84.61			58.33	60.07
$\tau = 0.2$		84.61			58.33	60.08
	96.44	83.08	71.56	59.31	56.03	60.76
	97.17	85.28			58.78	59.60
$\tau = 0.4$		85.33			58.83	59.53
	96.39	82.44	67.72	55.17	51.97	60.93
	97.28	85.39			59.06	59.55
$\tau = 0.6$		85.42			59.00	59.57
	96.25	81.50	62.17	51.06	47.92	61.87
	97.28	84.69			58.81	60.07
		84.64			58.67	60.09

4 结束语

本文针对野生动物监测图像物种分类任务中模型结果复核过程成本高昂的问题,提出了基于物种分类树的野生动物监测图像物种零样本层次化分类方法。实验结果表明,相比于路径矫正策略增强的基线方法,本文方法在纲、目、科、属和种5个分类层级上将准确率分别提高了0.27%、1.44%、1.42%、1.39%和1.03%,此外还使错误严重程度降低了1.1%。本文还探究了层次化分类中层级数量以及各层级分类间的独立程度对性能的影响,发现合理减少分类层级并约束层级间的影响程度有助于进一步提升层次化分类性能。

参考文献

- [1] RUSSAKOVSKY O, DENG J, SU H, et al. ImageNet large scale visual recognition challenge[J]. International Journal of Computer Vision, 2015, 115: 211-252.
- [2] CHEN G, HAN T X, HE Z, et al. Deep convolutional neural network based species recognition for wild animal monitoring [C]//Proceedings of IEEE International Conference on Image Processing (ICIP). Piscataway, NJ:IEEE, 2014: 858-862.
- [3] 宫一男, 谭孟雨, 王震, 等. 基于深度学习的红外相机动物影像人工智能识别:以东北虎豹国家公园为例[J]. 兽类学报, 2019, 39(4): 458-465.
- [4] CHU G, POTETZ B, WANG W, et al. Geo-aware networks for fine-grained recognition [C]// Proceedings of IEEE/CVF International Conference on Computer Vision Workshops (ICCVW). Piscataway, NJ:IEEE, 2019: 247-254.
- [5] NOROUZZADEH M S, NGUYEN A, KOSMALA M, et al. Automatically identifying, counting, and describing wild animals

- in camera-trap images with deep learning[J]. Proceedings of the National Academy of Sciences, 2018, 115(25): E5716-E5725.
- [6] YU H, XU Y, ZHANG J, et al. Ap-10k: A benchmark for animal pose estimation in the wild[J]. arXiv preprint arXiv, 2108.12617, 2021.
- [7] TUIA D, KELLENBERGER B, BEERY S, et al. Perspectives in machine learning for wildlife conservation [J]. Nature Communications, 2022, 13(1): 792.
- [8] BERTINETTO L, MUELLER R, TERTIKAS K, et al. Making better mistakes: Leveraging class hierarchies with deep networks [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ: IEEE, 2020: 12506-12515.
- [9] SILLA C N, FREITAS A A. A survey of hierarchical classification across different application domains [J]. Data Mining and Knowledge Discovery, 2011, 22: 31-72.
- [10] RADFORD A, KIM J W, HALLACY C, et al. Learning transferable visual models from natural language supervision[C]//Proceedings of International Conference on Machine Learning (ICML). IMLS, 2021: 8748-8763.
- [11] BABBAR R, PARTALAS I, GAUSSIER E, et al. On flat versus hierarchical classification in large - scale taxonomies [C]//Proceedings of Advances in Neural Information Processing Systems. Piscataway, NJ: IEEE, 2013: 26.
- [12] WU H, MERLER M, UCEDA-SOSA R, et al. Learning to make mistakes: Semantics-aware visual food recognition [C]//Proceedings of the 24th ACM International Conference on Multimedia. New York: ACM, 2016: 172-176.
- [13] WANG Y, WANG Z, HU Q, et al. Hierarchical semantic risk minimization for large-scale classification[J]. IEEE Transactions on Cybernetics, 2021, 52(9): 9546-9558.
- [14] YAN Z, ZHANG H, PIRAMUTHU R, et al. HD-CNN: Hierarchical deep convolutional neural networks for large scale visual recognition [C]//Proceedings of the IEEE International Conference on Computer Vision. Piscataway, NJ: IEEE, 2015: 2740-2748.
- [15] INOUE M, FORSTER C H, SANTOS A C. Semantic hierarchy-based convolutional neural networks for image classification[C]//Proceedings of International Joint Conference on Neural Networks (IJCNN). Piscataway, NJ: IEEE, 2020: 1-8.
- [16] ZHANG X, TANG L, LUO H, et al. Hierarchical bilinear convolutional neural network for image classification [J]. IET Computer Vision, 2021, 15(3): 197-207.
- [17] CHANG D, PANG K, ZHENG Y, et al. Your “Flamingo” is my “Bird”: fine-grained, or not[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ: IEEE, 2021: 11476-11485.
- [18] NOVACK Z, MCAULEY J, LIPTON Z C, et al. Chils: Zero-shot image classification with hierarchical label sets [C]//Proceedings of International Conference on Machine Learning (ICML). IMLS, 2023: 26342-26362.
- [19] REN Z, SU Y, LIU X. Chat GPT - powered hierarchical comparisons for image classification [C]// Proceedings of Advances in Neural Information Processing Systems (NeurIPS). 2024: 36.
- [20] KOH P W, SAGAWA S, MARKLUND H, et al. WILDS: A benchmark of in-the-wild distribution shifts[C]// Proceedings of International Conference on Machine Learning (ICML). IMLS, 2021: 5637-5664.