

融合多分支与多粒度特征的东北虎重识别

李晓楠¹, 朱 朦^{2,3}, 任洪娥^{1,3*}, 陶 锐^{1,4}

(1. 东北林业大学 信息与计算机工程学院, 哈尔滨 150040; 2. 哈尔滨学院, 哈尔滨 150086;
3. 黑龙江林业智能装备工程研究中心, 哈尔滨 150040; 4. 呼伦贝尔学院, 内蒙古 呼伦贝尔 021008)

摘 要:为解决东北虎重识别研究中存在的细节特征提取不充分等问题,提出了一种融合多分支与多粒度特征的东北虎重识别模型 CMM-Net。其中,全局分支负责提取宏观上的粗粒度特征;注意力分支通过插入坐标注意力模块加深了网络对重要特征的关注度;局部分支通过将特征图切分成不同条带块,从而提取东北虎更细粒度的局部特征。通过多个分支结构和多个细粒度特征结合来对模型进行优化学习,加强全局特征与局部特征的关联性。同时提出用 Circle Loss 与 Softmax 的联合损失来提高网络精度。实验结果表明,在 ATRW 数据集上所提模型在单摄像头环境下 mAP 为 93.6%,跨摄像头环境下 mAP 为 77.4%,均优于多数文献所提方法,证明了本文模型的有效性。

关键词:东北虎重识别;残差网络;多分支特征;坐标注意力机制;circle loss

DOI:10.15938/j.jhust.2023.06.011

中图分类号: TP751.1 **文献标志码:** A **文章编号:** 1007-2683(2023)06-0095-08

Re-Identification of Amur Tigers by Combining Multi-branch and Multi-granularity Features

LI Xiaonan¹, ZHU Meng^{2,3}, REN Honge^{1,3*}, TAO Rui^{1,4}

(1. College of Information and Computer Engineering, Northeast Forestry University, Harbin 150040, China;
2. Harbin University, Harbin 150086, China;
3. Heilongjiang Forestry Intelligent Equipment Engineering Research Center, Harbin 150040, China;
4. Hulunbuir University, Hulunbuir 021008, China)

Abstract: In order to solve the problem of insufficient detailed feature extraction in the re-identification of the Amur tiger, a re-identification model of the Amur tiger, CMM-NET, was proposed, which combined multi-branch and multi-granularity features. The global branch is responsible for extracting macroscopic coarse-grained features. The attention branch deepens the network's attention to important features by inserting coordinate attention module. Local branches can extract finer grained local features of amur tigers by cutting the feature map into different blocks. Finally, the model is optimized by combining multiple branch structures and multiple fine-grained features to strengthen the correlation between global features and local features. Meanwhile, the combined Loss of Circle Loss and Softmax is proposed to improve the network accuracy. Experimental results show that the mAP of the model proposed on THE ATRW data set is 93.6% in a single camera environment and 77.4% in a cross-camera environment, both of which are better than the methods proposed in most literatures, proving the effectiveness of the proposed model.

Keywords: re-identification of amur tiger; residual network; multiple branching feature; coordinate attention; circle loss

收稿日期: 2022-04-15

基金项目: 黑龙江省自然科学基金项目(LH2020F040);中央高校基本科研业务费专项资金资助项目(No. 2572017PZ10)。

作者简介: 李晓楠(1997—),女,硕士研究生;

朱 朦(1989—),女,博士,讲师。

通信作者: 任洪娥(1962—),女,教授,博士生导师,E-mail: nefu_rhe@163.com.

0 引 言

近几年,东北虎作为一种濒危的重点保护野生动物,对其进行有效的评估和监测已经成为相关研究人员关注的重要课题^[1-2]。由于虎群密度低、行为隐蔽、所处环境复杂等特点,使得其在检测过程中很难被发现。传统的红外触发相机陷阱技术^[2]需要专业人员进行人工分类和识别,耗费了大量的时间和精力。因此,使用深度学习的方法对东北虎进行自动追踪和分类有重要的研究意义。

重识别(Re-ID)是指检测图像或者视频中是否存在特定目标的计算机视觉技术,简单来讲就是给定一个监控目标,检索在其他监控设备下的该目标。早期的重识别技术主要针对行人和车辆进行研究,包括实现场景匹配以及目标的跨设备跟踪等,并且在目标跨境追踪和公共安全领域被广泛应用^[3]。传统的行人重识别需要手动提取复杂特征,且提取的特征有限。随着卷积神经网络在图像识别技术上的发展,基于深度学习的行人重识别逐渐成为研究热点^[4]。

基于深度学习的重识别一般分为特征提取和度量学习^[5-6]两部分。受深度学习在人脸识别项目中成功的启发,研究人员注意到可以用重识别技术获取野生动物的种群数量和行动轨迹^[7-8],从而实现对野生动物的保护。Carter 等人^[9]根据龟壳的图案研究了一种基于网络集成方法的海龟重识别系统;Weideman 等人^[10]将积分曲率表示与两种匹配算法结合,从而实现在鳍类中识别鲸类物种;Körschens 等人^[11]将目标部件定位与 CNN 特征和支持向量机分类相结合,设计了一个野外大象识别系统。近几年也有研究人员对东北虎进行了重识别研究,如 Li 等人^[12]介绍了一个东北虎重识别数据集,并提出了一种新的东北虎重识别方法,该方法通过对特征进行姿态局部对齐来处理东北虎较大的位姿变化,但由于此方法需要对东北虎的关键节点进行标注,容易因提取缺少或不正确的姿态信息而失败;Liu 等人^[13]提出了一种基于东北虎姿态特征的局部-全局引导框架-PPGNet 重识别网络,但此模型使用 ResNet-101 网络为基础,大大增加了模型的计算量;Yu^[14]等人采用一种基于局部分块和自适应 L2 正则化方法的东北虎重识别网络结构 PCB-AL2,其将

局部特征和全局特征结合起来,通过局部特征指导全局特征的学习,但此方法在跨摄像头等复杂环境下对东北虎进行重识别的精度还有待提高。

目前,针对动物进行分类检测的重识别算法较少,特别是针对东北虎进行重识别的研究都只用单一方法提取其局部特征,如用关键点和姿态等。而与行人特征相比,东北虎具有复杂的身体条纹信息和较大的活动姿态变化,所以基于结构信息的局部特征提取对于东北虎重识别效果不佳,而全局特征与多粒度特征结合,不仅可以从宏观上提取东北虎身体轮廓特征,还可以提取更细节的可区别于其他老虎的身体条纹特征,并且,在行人重识别研究观察中发现,随着分割粒度的增加,模型能学到更详细的细节信息。基于此,本文提出一种融合多分支与多粒度特征的东北虎重识别模型(Combining Multi-branch and Multi-granularity features, CMM-Net),该模型将东北虎多粒度细节特征与全局粗粒度特征结合,来弥补单一分支特征提取能力的不足。并且本文在损失函数上做进一步改进,用 Softmax 与 Circle Loss^[15]的联合损失来代替重识别中常用的三元组损失^[16],从而提升模型的精度。

1 融合多分支与多粒度特征的东北虎重识别模型

1.1 网络结构设计

本文提出的融合多分支与多粒度特征的东北虎重识别模型 CMM-Net 结构如图 1 所示,从左到右主要由骨干网络和三个分支网络构成。CMM-Net 的骨干网络选用轻量级 ResNet-50 网络,并对其进行修改,删除最后的池化层和全连接层,并在骨干网络 res_conv4_2 层之后分成 3 个分支模块,从上到下依次为全局分支(Global Branch)、注意力分支(Attention Branch)和局部分支(Local Branch),其中局部分支又分为 v1, v2 和 v3 3 个小分支,最后将所有经过卷积层降维后得到的 8 个分支特征串联,得到特征向量 F。为防止噪音干扰,整个网络只共享骨干网络的部分参数。

1.2 全局分支

为提取东北虎宏观的粗粒度特征,全局分支(Global Branch)在 res_conv5_1 层中使用 stride = 2 的卷积进行下采样,将骨干网络 res_conv5x 层输出

的 feature map 经卷积核为 4×12 的池化层之后生成全局分支特征 Z_g 。接着通过 1×1 卷积层进行批量归一化和 ReLU 操作,将特征 Z_g 从 2048-dim 降到 256-dim 得到特征 F_g 。即 Z_g 是 2048-dim 全局粗粒

度特征, F_g 是由 Z_g 降维后得到的 256-dim 特征。值得一提的是,为了保留更多图像信息,本文采用全局平均池化和全局最大池化的和作为骨干网络之后的池化层。

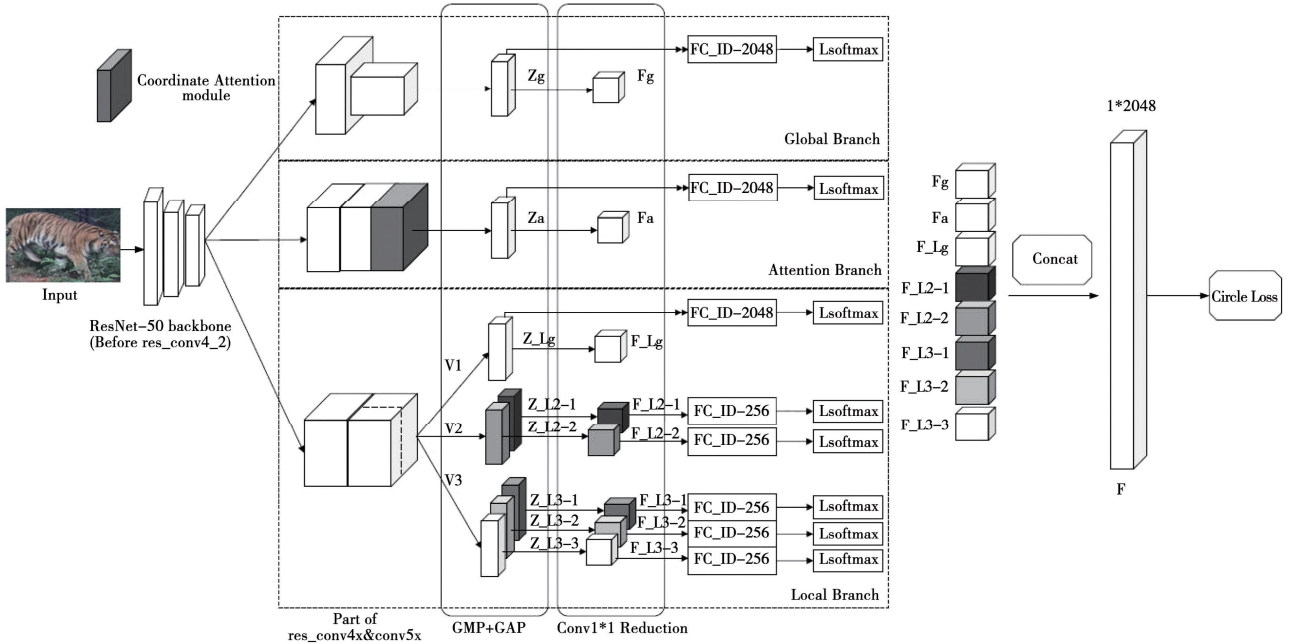


图1 CMM-Net 网络结构

Fig. 1 The structure of CMM-Net

1.3 注意力分支

注意力分支 (attention branch) 为了保留更多的东北虎特征信息,在骨干网络的 res_conv5_1 层中没有进行下采样操作,所以与全局分支输出的特征图相比尺寸大了一倍。在类似 ResNet-50 这种轻量级网络上的研究表明,通道注意力会给模型带来比较明显的性能提升效果,比如之前优秀的注意力模块 SENet^[17] 和 CBAM (convolutional block attention module)^[18]。但通道注意力通常会忽略特征图的位置信息,而这对于生成空间选择性注意映射有重要作用。其他能解决这个问题的注意力模块参数量太大,不适合应用于轻量级网络。因此,本文提出将 CA (coordinate attention, CA)^[19] 模块插入分支网络中,增强模型特征提取能力的同时也避免了网络产生大量的计算开销。该分支特征图经骨干网络 res_conv5x 层输出后进入 CA 模块,其结构如图 2 所示。经坐标注意力模块输出的特征图经过卷积核为 8×24 的池化层后得到注意力分支特征向量 Z_a ,再通过 1×1 卷积层最终得到降维特征向量 F_a ,大小为 256-dim。

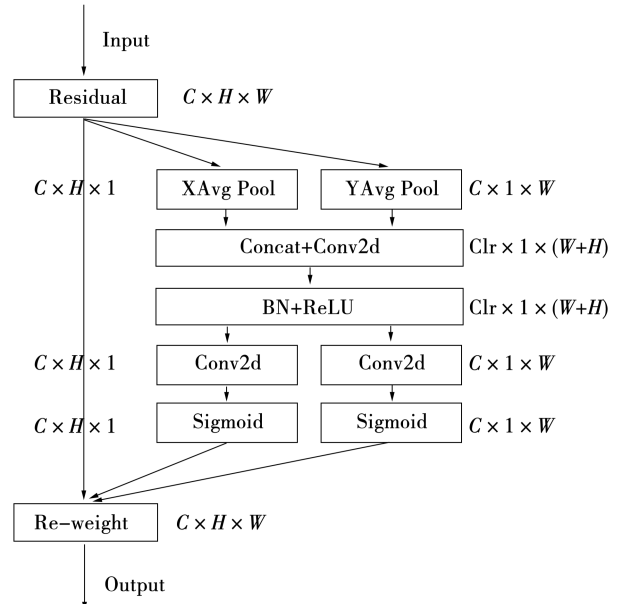


图2 坐标注意力模块结构

Fig. 2 The structure of Coordinate Attention module

CA 模块通过精准的位置信息对通道关系和长期依赖型进行编码,具体可分为坐标信息嵌入部分和坐标注意力生成部分。信息嵌入部分为了促进注

注意力模块能捕获具有精确位置信息的空间依赖,将全局池化分解为一对一维特征编码操作。对输入 x_c ,首先使用尺寸 $(H,1)$ 和 $(1,W)$ 的池化在水平坐标和垂直坐标两个方向分别进行通道编码,高度为 h 的第 c 个通道的输出表述如下:

$$z_c^h(h) = \frac{1}{W} \sum_{0 \leq i < W} x_c(h, i) \quad (1)$$

类似的宽度为 w 的第 c 个通道的输出表述如下:

$$z_c^w(w) = \frac{1}{H} \sum_{0 \leq j < H} x_c(j, w) \quad (2)$$

上述变换沿着空间的两个方向进行融合,返回一对方向感知注意力图。然后引入坐标注意力生成模块,首先级联之前模块生成的特征图,再用 1×1 卷积变换函数 F_1 ,得到 f :

$$f = \delta(F_1([z^h, z^w])) \quad (3)$$

其中 δ 是非线性激活函数, f 为在水平和垂直方向对空间信息进行编码的中间特征映射。再将 f 沿空间维度拆分为两个独立的张量 f^h 和 f^w , 并利用两个 1×1 卷积将特征图 f^h 和 f^w 变换到和输入 x_c 同样的通道数,得到如下结果:

$$g^h = \sigma(F_h(f^h)) \quad (4)$$

$$g^w = \sigma(F_w(f^w)) \quad (5)$$

其中 σ 是 Sigmoid 函数。输出 g^h 和 g^w 分别展开作为注意力权重。最后 CA block 的输出 y_c 可以写成:

$$y_c(i, j) = x_c(i, j) g_c^h(i) g_c^w(j) \quad (6)$$

到此, CA 模块同时完成了水平和垂直方向的注意力特征提取,并且能够有效地捕捉特征通道间的关系。

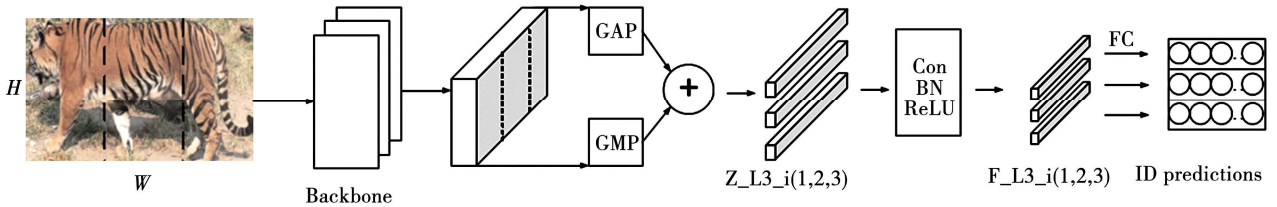


图 3 局部分支-V3 结构

Fig. 3 Local branch-V3 structure

1.4 局部分支

局部分支(local branch)在骨干网络 res_conv5_1 处同样没有进行下采样操作,且受 MGN^[20] 网络的启发,深度学习能自动关注整体图片中感兴趣的信息,图片分割后它也能在局部图片中关注到更丰富的细节信息。通过全局特征与多粒度局部特征结合,使得全局特征负责提取最粗粒度信息,而针对一些不显著且出现次数较少的细节特征,可通过分块数量的增加,每一块去负责不同层次或不同级别特征的提取。

因此针对东北虎难以辨别的条纹特征,本分支又划分为三个小分支 v1、v2 和 v3 分别对特征图的不同部位进行垂直切分,生成不同身体部位的细粒度特征块。其中, v1 分支使用卷积核为 8×24 的池化层将特征图分成一块,生成 $1 \times 1 \times 2048$ 的特征向量 Z_{L_g} ,之后通过 1×1 卷积进行降维得到 v1 分支特征向量 F_{L_g} ,大小为 256-dim。v2 分支与 v3 分支分别对特征图垂直划分,如图 3 所示为 v3 分支的结构,即将 $8 \times 24 \times 2048$ 的输出尺寸根据特征图的 W 轴分别划分为两个特征块和 3 个特征块,每个特征

块的大小为 $C \times H \times (W/n)$, n 分别为 2 和 3。并分别对切分后的各个特征块进行池化操作和卷积降维操作,最终 v2 分支生成 F_{L2-1} 和 F_{L2-2} 两个大小为 256-dim 的特征向量, v3 分支生成 F_{L3-1} 、 F_{L3-2} 和 F_{L3-3} 3 个大小为 256-dim 的特征向量。

1.5 损失函数

为了提高模型的表征学习能力,本文结合多分类任务 Softmax 与深度度量学习 Circle Loss 对模型提取的特征进行约束。为避免模型调参和收敛的困难,先对每个分支输出的特征单独进行分类,即对卷积前大小为 2048 维的全局特征 (Z_g 、 Z_a 和 Z_{L_g}) 和卷积后大小为 256 维的部分特征 (F_{L2-1} 、 F_{L2-2} 、 F_{L3-1} 、 F_{L3-2} 、 F_{L3-3}) 采用 Softmax 损失。针对 Circle Loss 本文采用文[15]的方法,将所有经过 1×1 卷积后的 256 维分支特征串联成单一特征向量 F ,再对 F 进行 Circle Loss。

传统的分类任务是将真实的标签看作 one-hot 的形式,但在训练数据较少的情况下,无法保证模型的泛化能力,容易造成过拟合。为解决这个问题本文采用标签平滑正则化方法(label smoothing),它可

以给 one-hot 标签做平滑处理,使其变成有概率值的软标签,即在真实处的标签概率值最大,而在其他位置的概率很小。这样标签平滑后的分布就相当于往真实分布中加入噪声,避免模型对于正确标签过于自信而造成过拟合,提高了模型的泛化能力。具体可表示为:

$$L_{\text{softmax}} = - \sum_{i=1}^N q_i \log(p_i) \quad (7)$$

$$q_i = \begin{cases} 1 - \frac{N-1}{N} \varepsilon, & i = \text{target} \\ \varepsilon/N, & i \neq \text{target} \end{cases} \quad (8)$$

其中: N 为类别总数; p_i 为预测为 i 的概率; q_i 为标签; ε 为较小的超参数。

Triple Loss 作为重识别常用的一种损失函数,它对于正负样本采用平均用力的方式进行优化,如图4所示。这样通常会使得模型在收敛时对于正负样本的区分力度不够,针对这一问题,Circle Loss 通过在正负样本对间重新加权,控制正负样本对各自梯度的贡献,使参数在训练过程中可以对类内相似性和类间相似性的反向传播梯度进行不同程度的扩大,以此来最大化类内相似性 S_p ,最小化类间相似性 S_n ,从而得到一个更有区分力度的模型。通过实验表明,Circle Loss 在重识别研究等多个视觉领域任务中取得了优势^[15]。

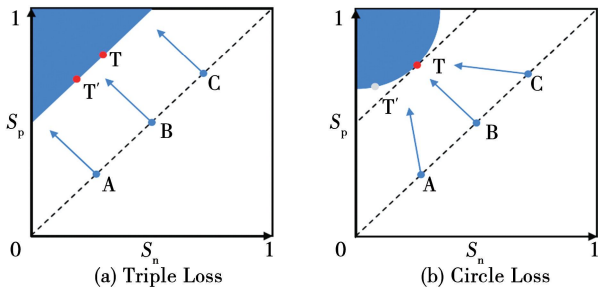


图4 空间维度

Fig. 4 Space dimension

值得一提的是,Circle Loss 对类别标签和样本对标签同时兼顾,其输出表述如下:

$$L_{\text{circle}} = \log \left[1 + \sum_{i=1}^K \sum_{j=1}^L \exp(\gamma(s_n^j - s_p^i + m)) \right] = \log \left[1 + \sum_{j=1}^L \exp(\gamma(s_n^j + m)) \sum_{i=1}^K \exp(\gamma(-s_p^i)) \right] \quad (9)$$

其中: K 为类内相似度得分; L 为类间相似度得分; γ 为尺度参数; m 为优化的严格程度。

2 实验与结果分析

2.1 数据集与数据处理

ATRW 数据集^[12]是在世界野生动物基金会(WWF)的帮助下,Maker Collider 从多个野生动物园的无约束自然环境中捕捉到的东北虎图像数据,是东北虎重识别的基准数据集。图5为该数据集的部分示例,a 为同一只老虎由于四肢不受限制而造成较大的姿态变化;b 和 c 分别为同一只老虎因在野生环境中被拍摄,所以有不同程度的遮挡和不受控制的光照条件;d 为条纹相似的不同老虎。此数据集共有 92 只老虎,并且以图像中老虎的视图方向(正面,左面,右面,背面)分别进行标注用来作为不同实体,其中一共包括 182 个实体的 3 649 个边界框注释。与行人重识别数据集不同的是,在这个数据集中并不是所有的实体都出现在跨摄像机中,约有 70 个实体来自跨摄像机,其余来自单个摄像机的不同帧。



(a) 不同姿势 (b) 不同遮挡程度 (c) 不同光照条件 (d) 条纹相似

图5 ATRW 数据集示例

Fig. 5 Examples of Amur tiger samples in the ATRW

在本文实验中,训练集包含 1 887 副图片,测试集包含 1 762 副图片。同时,因同一只东北虎的不同侧面被视为不同实体,为扩大数据集,本文将训练集的图像水平翻转来创建更多的新 ID,使训练集的图片数是原始图片数的两倍。并对输入网络的图片进行 3 种数据增强:从 -5° 到 5° 随机采样旋转角度;将亮度和对比度分别在 0.8 到 1.2 之间随机变化;图像随机仿射变换。

2.2 实验设置与评价指标

实验在 64 位 Ubuntu16.04 操作系统下,使用基于 Python 的深度学习框架 Pytorch 实现,并采用两块 NVIDIA RTX2080Ti 显卡进行 GPU 加速。本文使

用 ImageNet 中预训练的 ResNet-50 权重参数对骨干网络进行初始化。在训练阶段,将输入图像的大小调整为 128×384 ,模型共进行 220 次 epoch,每个训练批次大小为 32,为了抑制过拟合,使用 Adam 优化器代替 SGD 优化器,L2 正则化的权值衰减因子设置为 0.000 5,初始学习率为 2×10^{-4} ,标签平滑正则化损失函数的参数 ε 设置为 0.1,Circle Loss 的参数 γ 设置为 128, m 设置为 0.25。在推理阶段,只有全局分支提取的特征用于重识别,从而大大减少了计算量,因此本文提出的模型对野外东北虎重识别非常有效。

本实验采用两个评估指标,分别是平均精度均值(mean average precision, mAP)和累计匹配特征(cumulative match characteristics, CMC) Rank-1 精度。由于查询样本在相机中的出现情况,本实验将查询图像分为单摄像头和跨摄像头两个部分。

2.3 消融实验

为验证本文所提模型中各个分支的有效性,在 ATRW 数据集上根据分支的数量和类型分为 7 个实验进行评估,如表 1 所示。首先,分别进行 3 个单分支的实验;其次是任意两个分支互相组合的实验;最后是 3 个分支一起组合的实验。

表 1 不同分支的消融实验(%)

Tab. 1 Ablation experiments of different branches(%)

算法	单摄像头		跨摄像头	
	mAP	Rank-1	mAP	Rank-1
global	83.5	93.7	65.3	92.8
attention	85.7	94.8	67.9	93.6
local	89.2	96.2	72.4	94.5
global + attention	87.5	96.4	70.8	95.1
global + local	91.1	97.5	74.7	95.6
attention + local	91.6	97.8	75.2	96.2
global + attention + local	93.6	98.4	77.4	96.5

从表中可以看出,任意两个分支结合的识别精度都比单个分支识别精度高,并且 3 个分支结合时识别精度达到最高。即随着每个分支的加入,模型的精度都能得到不同程度提高,这说明每个分支都能提取到东北虎的不同特征,各个分支提取的细粒度特征与全局粗粒度特征互补,相互监督学习,从而提升了模型的精度。

2.4 与现有方法比较

为进一步验证本文所提模型 CMM-Net 的有效性,选取部分东北虎重识别文献方法与本文方法作比较,结果如表 2 所示。

表 2 在 ATRW 数据集上与文献方法比较(%)

Tab. 2 Compared with the literature method on ATRW data set(%)

算法	单摄像头		跨摄像头	
	mAP	Rank-1	mAP	Rank-1
PPbM ^[12]	72.8	89.4	47.8	77.1
PPGNet ^[13]	90.6	97.7	72.6	93.6
PCB-AL2 ^[14]	91.5	97.8	74.5	94.9
CMM-Net(ours)	93.6	98.4	77.4	96.5

从表中可以看出本文提出的模型在单摄像头和跨摄像头环境下的识别精度均优于文献方法,在单摄像头环境下 mAP 达到 93.6%,Rank-1 达到 98.4%,与 PCB-AL2 相比性能分别提高了 2.1% 和 0.6%。在跨摄像头环境下 mAP 达到 77.4%,Rank-1 达到 96.5%,与 PCB-AL2 相比性能分别提高了 2.9% 和 1.6%。值得一提的是,在相同实验参数下,用文[13]、[14]方法测试查询集的重识别速度分别为 12.8 mins 和 5.6 mins,本文方法为 8.20 mins,可以看出虽然仅比最短时间多了 2.6 mins,但最终精度得到明显提高,证明了本文模型 CMM-Net 对东北虎重识别研究的有效性。同时,从表 2 可以看出在 ATRW 数据集上所有文献方法在单摄像头环境下的性能指标都要优于跨摄像头,这是因为单摄像头采集的数据都是东北虎在某个摄像头下被连续拍到的图像,其背景环境,光照条件等方面的差距较小,而跨摄像头采集的是东北虎被两个或者两个以上摄像头拍摄的图像,其动作范围较广,姿态变化、遮挡程度和背景环境的差距较大,因此对跨摄像头环境下的东北虎进行分类和追踪有重要研究意义。

2.5 可视化实验

为进一步验证本文模型的有效性,本节将 CMM-Net 模型和只含全局分支的结构模型作为基线进行可视化对比,即使用 CAM 实现特征提取过程中 layer4 的可视化热力图,其部分可视化结果如图 6 所示,图中区域颜色越亮,即训练过程中模型关注的越多。

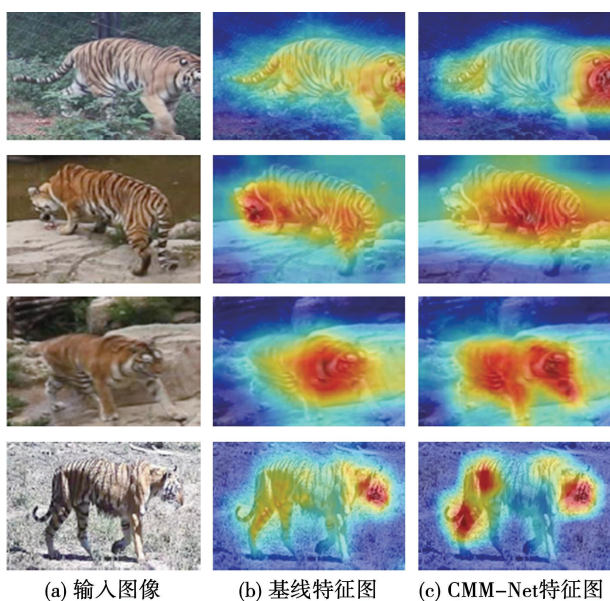


图6 可视化特征图

Fig. 6 Visual features

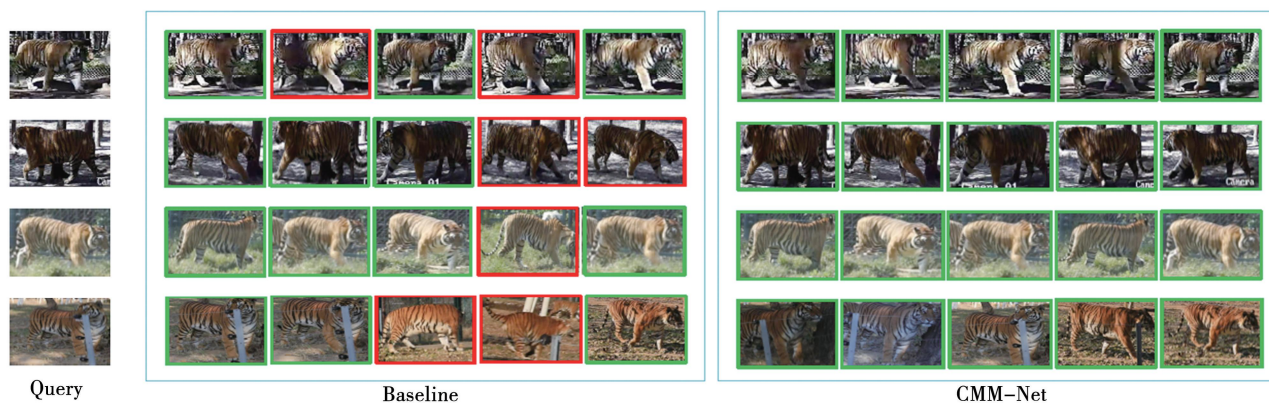


图7 ATRW 数据集部分图像查询结果

Fig. 7 ATRW dataset partial image query results

3 结论

本文提出了一种融合多分支与多粒度特征的东北虎重识别模型 CMM-Net, 通过将全局分支、注意力分支与局部分支提取的特征进行联合学习, 使模型在训练阶段两部分细粒度局部分支特征都可以作为一种调节器来约束粗粒度全局特征的学习, 从而弥补单一分支特征提取能力的不足。并且相较于重识别研究中常用的交叉熵与三元组损失, 本文采用标签平滑正则化 Softmax 函数与 Circle Loss 对模型进行收敛训练, 提高网络精度。实验结果表明, 本文模型 CMM-Net 在 ATRW 数据集上的 mAP 在单摄像头和跨摄像头环境下分别达到 93.6% 和 77.4%, 能

由图可见, 基线的回应主要集中在身体的主干部分, 而头部, 脚部等细节信息容易被忽略, 而 CMM-Net 模型高激活区域分布的范围更集中, 更多地针对某些关节部位, 而这些部位的细粒度特征通常对辨别不同东北虎的身份起着关键作用, 能提取到更具有辨别力的东北虎特征进行重识别, 证明了本文模型将全局特征和细粒度局部特征结合的有效性。

此外, 本文还展示了在 ATRW 数据集下部分查询图像的前 5 个排序结果, 如图 7 所示。从图中的前两张图片可以看出, 当野外光照条件不充足时, 相较于基线, CMM-Net 能通过注意力机制发现更多具有辨别力的特征, 通过第三张和第四张图片可以看出, 针对不同身份却条纹相似的东北虎, CMM-Net 能通过提取细粒度局部信息进行正确区分, 即使在野外有遮挡的条件下, CMM-Net 通过全局与局部细节特征的结合, 从而使东北虎特征表示更具有鲁棒性。

够有效提升东北虎重识别精度。下一步我们将重点研究在跨域情况下如何提升东北虎重识别模型的泛化能力, 提取更具有判别性和鲁棒性的东北虎特征。

参考文献:

- [1] 顾佳音, 刘辉, 姜广顺. 东北虎 (*Panthera tigris altaica*) 个体识别技术研究进展 [J]. 野生动物, 2013, 34 (4): 229.
GU J, LIU H, JIANG G S. A Review of Potential Techniques for Identifying Individual Amur Tigers (*Panthera tigris altaica*) [J]. Chinese Journal of Wildlife, 2013, 34 (4): 229.
- [2] GOMEZ A, SALAZAR A, VARGAS-BONILLA JF. Towards Automatic Wild Animal Monitoring: Identification of Animal Species in Camera-trap Images Using Very

- Deep Convolutional Neural Networks[J]. *Ecological Informatics*, 2016, 41.
- [3] YE M, SHEN J, LIN G, et al. Deep Learning for Person Re-identification: A Survey and Outlook[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021(99):1.
- [4] 张珂,侯捷. 基于改进的卷积神经网络图像识别方法[J]. *科学技术与工程*, 2020, 20(1):6.
ZHANG K, HOU J. Image Recognition Method Based on Improved Convolutional Neural Network [J]. *Science Technology and Engineering*, 2020, 20(1):6.
- [5] SUN Y, WANG J, TANG S, et al. Part-aligned Bili-near Representations for Person Re-identification [C]//Proceedings of the European Conference on Computer Vision (ECCV), 2018:402.
- [6] ZHANG H, GOODFELLOW I, METAXAS D, et al. Self-attention Generative Adversarial Networks[C]//International Conference on Machine Learning. PML-R, 2019:7354.
- [7] BOUMA S, PAWLEY M, HUPMAN K, et al. Individual Common Dolphin Identification Via Metric Embedding Learning[C]//2018 International Conference on Image and Vision Computing New Zealand (IVCNZ), 2018.
- [8] HUGHES B, BURCHARDT T. Automated Visual Fi-n Identification of Individual Great White Sharks[J]. *International Journal of Computer Vision*, 2017, 122(3):542.
- [9] CARTER S, BELL I P, MILLER JJ, et al. Automated Marine Turtle Photograph Identification Using Artificial Neural Networks, with Application to Green Turtles[J]. *Journal of Experimental Marine Biology & Ecology*, 2014, 452:105.
- [10] WEIDEMAN H J, JABLONS Z M, HOLMBERG J, et al. Integral Curvature Representation and Matchi-ng Algorithms for Identification of Dolphins and W-hales[C]//Proceedings of the IEEE International Conference on Computer Vision Workshops, 2017: 2831.
- [11] Körschens M, BARZ B, DENZLER J. Towards Automatic Identification of Elephants in the Wild[J]. ar-Xiv preprint arXiv:1812.04418, 2018.
- [12] LI S, LI J, TANG H, et al. ATRW: A Benchmark for Amur Tiger Re-identification in the Wild[C]//Proceedings of the 28th ACM International Conference on Multimedia. 2020: 2590.
- [13] LIU C, ZHANG R, GUO L. Part-pose Guided AmurTiger Re-identification[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops. 2019.
- [14] 于慧伶,钱成帅. 基于自适应正则化的东北虎重识别方法[J]. *计算机工程与应用*, 2022, 58(8): 191.
YU H L, QIAN C S. Re-identification Method of S-iberian Tiger Based on Adaptive Regularization[J]. *Computer Engineering and Applications*, 2022, 58(8): 191.
- [15] SUN Y, CHENG C, ZHANG Y, et al. Circle loss: A Unified Perspective of Pair Similarity Optimization [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 6398.
- [16] SCHROFF F, KALENICHENKO D, PHILBIN J. Facenet: A Unified Embedding for Face Recognition and Clustering[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015: 815.
- [17] HU J, SHEN L, SUN G. Squeeze-and-Excitation Networks [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 7132.
- [18] WOO S, PARK J, LEE J Y, et al. Cbam: Convoluti-onal Block Attention Module[C]//Proceedings of the European Conference on Computer Vision (ECCV), 2018: 3.
- [19] HOU Q, ZHOU D, FENG J. Coordinate Attention for Efficient Mobile Network Design [J]. arXiv. 2013. 02907, 2021.
- [20] WANG G, YUAN Y, CHEN X, et al. Learning Discriminative Features with Multiple Granularities for Person Re-identification[C]//Proceedings of the 26th ACM International Conference on Multimedia, 2018: 274.

(编辑:温泽宇)