doi：10.13413/j.cnki.jdxblxb.2023431

基于多粒度融合和双注意力的细粒度图像分类

李鹏松[1]，周冰倩[1]，季芷伊[1]，于永平[2]
（1. 东北电力大学 理学院，吉林 吉林 132012；2. 吉林大学 建设工程学院，长春 130021）

摘要：针对现有模型对细粒度图像关键信息精准识别较难，分类指标较单一且特征利用不充分的问题，提出一个新的细粒度图像分类网络模型. 该模型在网络训练步骤中嵌入双注意力网络以强化中层特征与深度特征的相关性，根据网络不同层的感受野大小不同将数据剪裁后再拼接成新的样本数据作为下一层输入，采用支持向量机分类器将中层和深度特征输出结果一同作为最终分类指标. 在 3 个经典数据集 CUB-200-2011、Stanford Cars 和 102 Category Flower 上的实验结果表明，其分类准确率分别达 89.56%，95.00%，96.05%，相比于其他网络模型有较好的分类准确率和泛化能力.

关键词：细粒度图像分类；注意力机制；数据增强；多粒度特征融合
中图分类号：TP391.41 **文献标志码**：A **文章编号**：1671-5489(2024)06-1447-08

Fine-Grained Image Classification Based on Multi Granularity Fusion and Dual Attention

LI Pengsong[1], ZHOU Bingqian[1], JI Zhiyi[1], YU Yongping[2]
(1. *School of Science, Northeast Electric Power University, Jilin* 132012, *Jilin Province, China*;
2. *College of Construction Engineering, Jilin University, Changchun* 130021, *China*)

Abstract: Aiming at the problems that it was difficult to accurately identify the key information of fine-grained images, the classification index was relatively simple and the feature utilization was not sufficient in existing models, we proposed a new fine-grained image classification network model. In the network training step, the model embedded a dual attention network to strengthen the correlation between middle-level features and depth features. According to the different receptive field sizes of different layers of the network, the data were trimmed and then spliced into new sample data as the input for the next layer. The support vector machine classifier was used to take the output results of middle-level features and depth features together as the final classification index. The experimental results on three classic datasets CUB-200-2011, Stanford Cars and 102 Category Flower show that the classification accuracy reaches 89.56%, 95.00% and 96.05%, respectively. Compared with other network models, it has better classification accuracy and generalization ability.

Keywords: fine-grained image classification; attention mechanism; data augmentation; multi granularity feature fusion

区别于传统的图像分类任务，细粒度图像分类(fine-grained visual classification，FGVC)是一种将

收稿日期：2023-10-13.

第一作者简介：李鹏松(1970—)，男，汉族，博士，教授，从事非线性动力系统和深度学习的研究，E-mail：li-pengsong@163.com.
通信作者简介：周冰倩(1997—)，女，汉族，硕士研究生，从事人工智能理论及应用的研究，E-mail：1099480882@qq.com.

基金项目：国家自然科学基金(批准号：42372356).

图像进行更精细的分类工作，需要学习同一类别图像的较深层特征并做比较. 由于细粒度图像数据样本集中，同一类别下的基础类别间具有非常高的相似性，因此需要细粒度图像分类技术能精准地识别子类别间的细微差异.

在细粒度图像分类的许多方法中已证明局部检测对细粒度图像分类的性能有较大改善[1-4]. 一些方法[5-8]通过使用额外的标签帮助卷积神经网络在学习过程中捕捉同一子类的区分性局部特征. Farrell 等[9]提出了用于细粒度分类的姿态归一化表示法，尽管该方法取得了较好的分类效果，但需要人工标注，工作成本较高. 区别于局部特征提取，端到端的特征编码更倾向于增强图像信息的全局特征，进而对具有判别性全局特征进行学习，目前主流的端到端图像特征编码框架是双线性模型[10-13]，对细粒度图像分类有较好性能，但该方法一般提取特征尺度较单一，会减弱细粒度图像数据的位置信息.

近年来，一种可以精准关注图形有效区域的方法——注意力机制，被用于细粒度图像识别的局部信息特征提取中，其数学思想就是给局部关键性的信息特征分配较高的权重. 文献[14]首次将注意力机制应用到计算机视觉领域，使注意力机制在图像处理中取得了良好的效果. 朱丽等[15]设计了一种引入双路通道注意力的残差网络模型，并应用了多头自注意力机制，但该方法仅考虑了特征通道之间的关系，未考虑特征空间对细粒度图像分类的影响. Liu 等[16]提出了一种基于强化学习的全卷积注意力定位网络，采用全卷积网络架构可以避免全连接层的超大计算量，但局部区域的高分辨率处理会导致在多步前向和后向传播过程中很耗时. 在细粒度图像分类中，常见的方法大多数是对图像数据先捕捉关键性局部特征，再进行特征学习，这种方法虽取得了较好的效果，但都只关注了数据的局部信息特征，而很少考虑到注意力机制对不同粒度图像数据的影响，并且大多数方法均只考虑了深度特征的输出，将中层特征和深度特征综合考虑到最终结果中的研究报道较少.

针对上述问题，本文提出一种基于卷积块注意力模型(convolutional block attention module，CBAM)和多粒度融合(multi granularity fusion)的细粒度图像分类网络框架. 首先，在训练过程中嵌入卷积块注意力网络训练网络更高层，该注意力网络考虑了特征通道和特征空间两个维度的权重系数，有助于弱监督学习过程中更好地突出细粒度图像类内区分性特征，并考虑了注意力模块对不同粒度特征的影响；其次，在网络的前一阶段将样本数据特征图分割成不同小块，再应用拼图生成器代码生成新的特征图作为下一阶段的输入，使其可充分利用细粒度图像数据的特征，起到数据增强的作用；最后，在分类阶段，本文将特征图的中层和深度多粒度特征一同输出，用于损失计算和参数更新，以保证分类结果的全面性. 实验结果表明，本文提出的网络框架提升了数据集测试结果的准确率.

1 模型原理与架构

1.1 卷积块注意力网络模型

卷积块注意力模型如图1所示. 其是一种结合了空间和通道的注意力机制模块，相比于只关注通道或空间的注意力机制，它的效果更好.

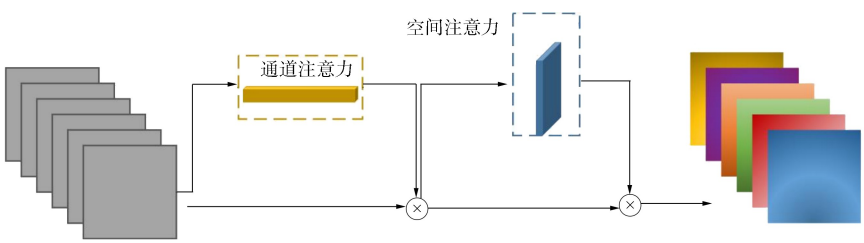


图1 卷积块注意力模型

Fig. 1 Convolutional block attention model

由图1可见，CBAM 在输入端和输出端之间分别将通道和空间注意力机制两个模块部分串联，按顺序对图像特征的通道和空间两个维度进行学习并计算得到注意力权重，将模型学习后得到的注意力权重与输入特征图矩阵相乘，得到细粒度图像显著特征图，计算公式如下：

$$\boldsymbol{F}''=\boldsymbol{M}_S(\boldsymbol{M}_C(\boldsymbol{F})\otimes\boldsymbol{F})\otimes(\boldsymbol{M}_C(\boldsymbol{F})\otimes\boldsymbol{F}),\tag{1}$$

其中 $\boldsymbol{F}\in\mathbb{R}^{C\times H\times W}$ 表示上一层的输入图像特征，$\boldsymbol{M}_C\in\mathbb{R}^{C\times1\times1}$ 和 $\boldsymbol{M}_S\in\mathbb{R}^{1\times H\times W}$ 分别表示卷积模块注意力中依次得到的通道注意力特征和空间注意力特征，$\boldsymbol{F}''$ 为得到的最终输出，$\otimes$ 表示特征相乘操作.

通道注意力模块[17]的构造是通过不同的压缩方式将平面压缩成一个值后，再经过全连接层连接后对特征进行重新组合并求积运算得到输出特征，如图 2 所示.

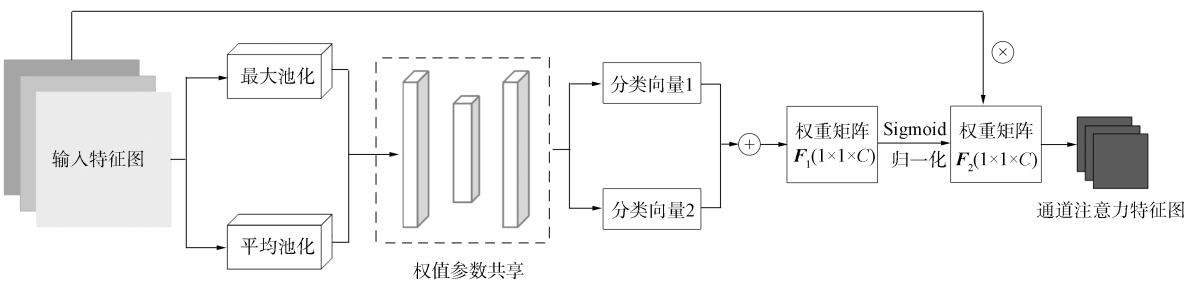


图 2　通道注意力网络模型

Fig. 2　Channel attention network model

由图 2 可见，分别对每个特征图 $\boldsymbol{F}\in\mathbb{R}^{H\times W\times C}$ 进行全局平均和最大池化操作以降低特征图维度. 先将池化后的特征图输入多层感知机中进行学习，再将两张特征图在通道维度上进行堆叠操作，经过 Sigmoid 激活函数层将特征图中每个通道的权重进行归一化，生成注意力权重，最后将归一化注意力权重与特征图相乘，生成空间注意力模块需要的输入特征为

$$\boldsymbol{M}_C(\boldsymbol{F})=\boldsymbol{F}\otimes\sigma(\mathrm{MLP}(\boldsymbol{F}^c_{\mathrm{avg}})\oplus\mathrm{MLP}(\boldsymbol{F}^c_{\mathrm{max}})),\tag{2}$$

其中 $\oplus$ 表示特征堆叠操作，$\otimes$ 表示特征相乘操作，σ 表示 Sigmoid 函数变换，$\boldsymbol{F}^c_{\mathrm{avg}}$ 和 $\boldsymbol{F}^c_{\mathrm{max}}$ 分别表示对输入特征图在通道维度上的全局平均池化和最大池化后的特征图，MLP(multi-layer perceptron)表示多层感知机操作.

通道注意力模块主要集中在给定输入图像“什么”信息是重要的，空间注意力模块[18]则关注重要信息区域“在哪里”，可以作为通道注意力的补充，如图 3 所示.

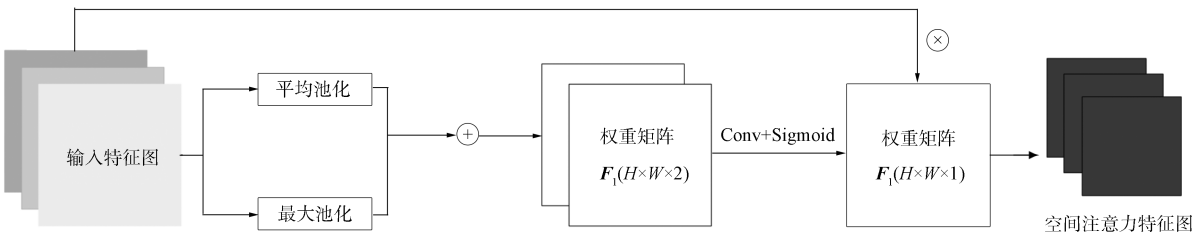


图 3　空间注意力网络模型

Fig. 3　Spatial attention network model

由图 3 可见，首先将通道注意力的计算输出特征作为本模块的输入，然后在同一通道下的不同空间对图像特征分别做池化计算，对池化后的两张特征图进行通道维度堆叠，再做卷积计算，将特征维度降为 1，以方便下一步进行的 Sigmoid 归一化计算，最后将输入特征图与归一化权重相乘，用公式表示为

$$\boldsymbol{M}_S(\boldsymbol{F})=\boldsymbol{F}\otimes\sigma(\boldsymbol{f}^{7\times7}*(\boldsymbol{F}^s_{\mathrm{avg}}\oplus\boldsymbol{F}^s_{\mathrm{max}})),\tag{3}$$

其中 $\boldsymbol{F}^s_{\mathrm{avg}}\in\mathbb{R}^{1\times H\times W}$ 和 $\boldsymbol{F}^s_{\mathrm{max}}\in\mathbb{R}^{1\times H\times W}$ 分别表示平均池化和最大池化后的特征图，$\boldsymbol{f}^{7\times7}$ 表示一个 7×7 卷积核，$*$ 表示卷积操作.

1.2　本文网络模型架构

本文改进模型使用 ResNet-50 网络作为基本特征提取网络，其原始框架如图 4 所示. 由图 4 可见，该网络架构主要部分包括 1 个卷积模块、4 个残差模块和 1 个输出层.

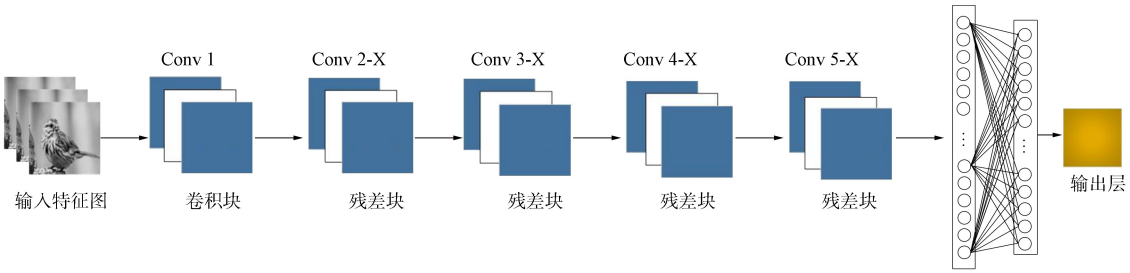


图 4 原模型结构

Fig. 4 Original model structure

残差网络结构突破了传统对网络简单叠加的思想，从而实现了数据在模块内部特征的传递不会丢失，ResNet-50 网络叠加卷积和残差模块列于表 1.

表 1 ResNet-50 网络叠加卷积和残差模块

Table 1 Stacked convolution and residual module of ResNet-50 network

网络阶段	卷积块 1	残差块 2	残差块 3	残差块 4	残差块 5
输出尺寸	112×112	56×56	28×28	14×14	7×7
ResNet-50	3×3，最大池化，步长 2 7×7，64，步长 2	$\begin{pmatrix}1\times1,64\\3\times3,64\\1\times1,256\end{pmatrix}\times3$	$\begin{pmatrix}1\times1,128\\3\times3,128\\1\times1,512\end{pmatrix}\times4$	$\begin{pmatrix}1\times1,256\\3\times3,256\\1\times1,1\,024\end{pmatrix}\times6$	$\begin{pmatrix}1\times1,512\\3\times3,512\\1\times1,2\,048\end{pmatrix}\times3$

本文改进模型总体框架如图 5 所示. 由图 5 可见，该模型框架由基于双注意力机制的特征提取网络模块组成，将通道和空间特征信息融合到中层特征提取网络中，中层特征提取部分在原始网络中进行修改，并融合多粒度特征，同时对网络的输入图像数据进行再次拆分重拼接，保证数据充分被利用并提升网络模型的鲁棒性和泛化能力. 原始网络的分类方法只把最后一层的输出作为指标计算分类结果，本文将中层特征输出加入到最终分类指标中.

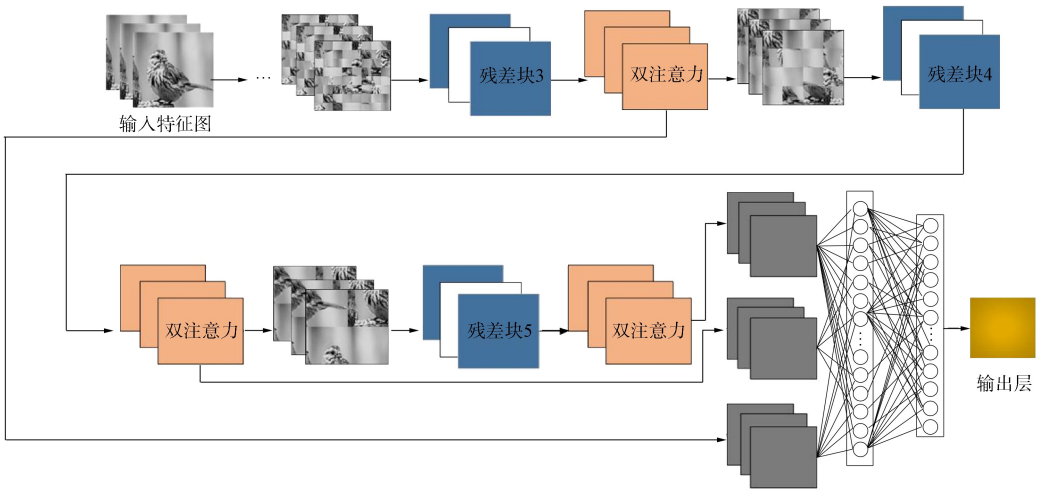


图 5 改进模型整体结构

Fig. 5 Overall structure of improved model

在网络经过两阶段卷积操作后，本文使用拼图生成器代码将特征图像分解后重组，对细粒度图像数据特征充分利用. 为强化局部关键性特征提取加入卷积注意力模块，对网络的深中层特征进行输出，将多粒度特征融合，再经过分类模块得到预测分类结果. 其中分类模块包括两个全连接层，使用 Softmax 激活函数将一个数值向量归一化为另一个实数向量，使新的实数向量中每个元素取值都介于(0,1)内，Softmax 函数计算公式如下：

$$\tilde{y}_{ij}=\frac{e^{\alpha_j}}{\sum_{j=1}^{k}e^{\alpha_j}},\tag{4}$$

其中 α_j 为输出层的第 j 个输出值，k 为数据集分类样本类别数，$\tilde{y}_{ij}$ 表示第 i 个样本数据为数据集中第 j 类的概率值. 本文算法使用随机梯度下降法不断迭代更新权重优化改进网络模型，权重系数更新公式如下：

$$\theta_{i+1}=\theta_i-\alpha\frac{\partial\sum_{i=1}^{n}f(\theta_1,\theta_2,\cdots,\theta_i)}{\partial\theta_i},\tag{5}$$

其中：θ_i 为第 i 次迭代的待求解权重系数；α 为下降系数，即学习率；$\frac{\partial\sum_{i=1}^{n}f(\theta_1,\theta_2,\cdots,\theta_i)}{\partial\theta_i}$ 为损失函数 $f(\theta_1,\theta_2,\cdots,\theta_i)$ 对第 i 次权重系数的导数.

2　实验与分析

2.1　实验设置与数据集

本文实验采用的服务器硬件配置为 Tesla T4 16 GB 的显卡，内存为 32 GB，CUDA11.4 的驱动，语言环境为 Python3.7，深度学习框架为 Pytorch1.7.1. 输入图像尺寸为 500×500，在训练阶段将其固定剪裁为 448×448 的尺寸，训练迭代次数为 300 次，batch_size 设为 16.

本文使用 3 个经典的细粒度图像分类数据集，分别为 CUB-200-2011(CUB)、Stanford Cars(Cars)和 102 Category Flower(Flower)，各数据集类别信息及划分情况列于表 2.

表 2　各数据集类别信息及划分情况

Table 2　Category information and classification of each dataset

数据集	类别数	图片数量/张	训练集 & 验证集数据量/张	测试集数据量/张
CUB-200-2011	200	11 788	5 994	5 794
Stanford Cars	196	16 185	8 144	8 041
102 Category Flower	102	8 189	2 040	6 149

2.2　实验结果与分析

本文构建了双注意力和多粒度特征融合的改进细粒度图像分类算法，并通过将训练图像数据拆分重组进而丰富了数据的多样性，提高了网络模型的泛化性能. 在鸟类、车类和花类细粒度 3 种公开数据集上实验，分类精确度较好. 为进一步验证本文算法中每个模块的有效性，进行了消融实验. 图 6 为消融实验与文本算法的对比结果. 结果表明，在 3 个数据集上准确率分别提升 0.74，0.2，0.39 个百分点.

由图 6(A)可见，数据集 CUB-200-2011 在优化迭代 55 次时，测试集准确率达 88.87%，最后经过 165 次迭代训练，算法损失收敛至 0.5 附近，准确率最高达 89.56%. 由图 6(B)可见，数据集 Stanford Car 优化迭代到 50 次时，测试集准确率达 94.6%，最后经过 170 次迭代训练，算法损失收敛至 0.4 左右，准确率最高达 95%. 由图 6(C)可见，数据集 102 Category Flower 在优化迭代 45 次时，测试集准确率达 94.56%，最后经过 115 次迭代训练后，准确率达最高为 96.05%. 由图 6(D)，(E)和(F)可见，本文改进网络模型在迭代约 100 次时收敛，损失值几乎不变，而不加卷积注意力模块的网络模型收敛较慢，并且本文模型在准确率提高的同时也更稳定，进一步验证了本文改进模型既可识别细粒度图像关键性区域，又可以捕捉中间层和深层抽象特征相关性和融合多种粒度信息特征，具有较好的分类性能. 因此，本文网络模型适用于图像背景复杂、类内差距不明显的细粒度图像分类问题.

为证明本文改进算法具有较优的分类准确率和泛化能力，表 3 列出了不同网络模型在 3 个数据集上的分类准确率对比结果.

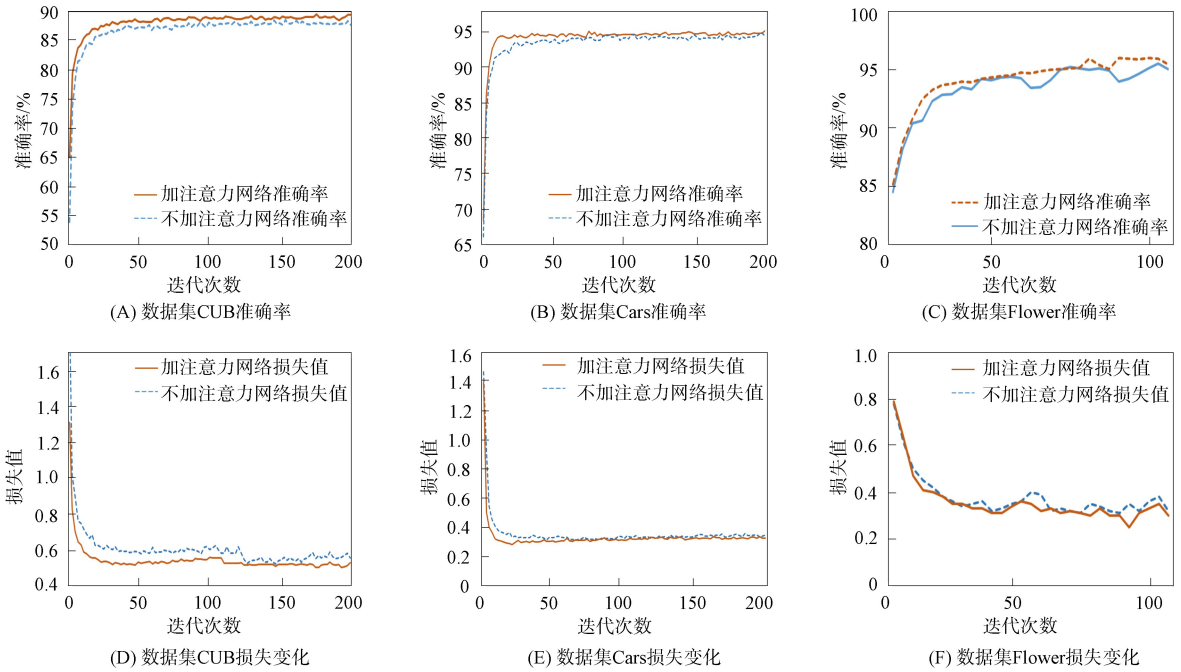


图 6 3个数据集中准确率及损失值消融实验结果

Fig. 6 Ablation experiment results of accuracy and loss values in three datasets

表 3 不同网络模型在3个数据集上的分类准确率对比结果

Table 3 Comparison results of classification accuracy of different network models on three datasets

网络模型	准确率/%		
	CUB-200-2011	Stanford Cars	102 Category Flower
DFL-CNN[19]	87.40	93.80	—
PCA-Net[20]	88.90	94.30	—
RA-CNN[21]	85.30	92.50	—
CABMN[22]	90.07	91.95	—
Gogul[23]	—	—	93.41
Zhu[15]	—	—	94.42
本文模型	89.56	95.00	96.05

在细粒度图像分类任务中，表3中的几种网络都是基于弱监督学习网络，DFL-CNN增强了中层学习能力，PCA-Net,RA-CNN,CABMN均是一种引入注意力机制的网络模型，本文的网络模型在数据集CUB-200-2011和Stanford Cars上分类准确率分别达89.56%和95.00%，与其他几种模型相比，鸟类数据集分类准确率分别提升了2.16,0.66,4.26个百分点，分类准确率略低于CABMN模型，车类数据集分类准确率分别提升了1.2,0.7,2.5,3.05个百分点. 文献[23]是使用CNN作为特征提取器并引入迁移学习的网络模型，文献[15]是一种引入双路通道注意力机制的网络模型，与两种网络模型相比，本文模型在数据集102 Category Flower上达到96.06%的分类准确率，准确率分别提升了2.64,1.63个百分点. 由表3可见，与其他的细粒度图像分类模型相比，本文网络模型的分类准确率较高，从而验证了该模型具有较好的分类效果.

实验结果表明，相比于其他的细粒度图像分类网络模型，本文构建的基于多粒度融合和双注意力的网络模型不仅从多维度对数据显著性特征进行学习，且保证了分类数据结果的全面性，通过在3个经典数据集上进行实验对比，网络具有较高的识别准确率，说明模型在细粒度图像分类上能更好地捕捉数据的显著特征，有较好的分类性能和一定的实际应用价值.

综上所述，针对现有模型对细粒度图像关键信息精准识别较难，分类指标较单一且特征利用不充分的问题，本文提出了一个多粒度融合和双注意力的细粒度图像分类网络模型. 该模型通过提取网络不同层特征进行多粒度特征融合起到数据增强作用，更好地定位了细粒度图像数据区分性特征.

此外，通过构建基于双注意力机制的特征提取网络，使模型能更好地学习不同空间和通道间的联系. 该算法的实现过程只用到数据类别标签的监督信息，模型可达到端到端训练. 最后在 3 个经典公开数据集上测试，结果表明，本文模型达到了较高的分类率，并且优于其他对比模型.

参 考 文 献

[1] ARDHENDU B，ZACHARY W. Context-Aware Attentional Pooling（CAP）for Fine-Grained Visual Classification [J]. Proceedings of the Conference on Artificial Intelligence，2021，35(2)：35-37.

[2] CHEN J M，HU J G，LI S R. Learning to Locate for Fine-Grained Image Recognition [J]. Computer Vision and Image Understanding，2021，206(3)：103-184.

[3] YANG S K，LIU S，YANG C，et al. Re-rank Coarse Classification with Local Region Enhanced Features for Fine-Grained Image Recognition [EB/OL]. (2021-02-19)[2023-09-06]. https://arxiv.org/abs/2102.09875.

[4] 张志林，李玉鑑，刘兆英. 深度学习在细粒度图像识别中的应用综述 [J]. 北京工业大学学报，2021，47(8)：942-953. (ZHANG Z L，LI Y J，LIU Z Y. A Review of the Application of Deep Learning in Fine Grained Image Recognition [J]. Journal of Beijing University of Technology，2021，47(8)：942-953.)

[5] AMOUS F I，NARAYANAN R M，LI B C. Application of Multidomain Sensor Image Fusion and Training Data Augmentation for Enhanced CNN Image Classification [J]. Journal of Electronic Imaging，2022，31(1)：013014-1-013014-8.

[6] LI H H，WEI Y N，ZHANG H M，et al. Fine-Grained Classification of Grape Leaves via a Pyramid Residual Convolution Neural Network [J]. International Journal of Agricultural and Bioengineering，2022，15(2)：197-203.

[7] ZHU H G，GAO Z，WANG J Y，et al. Few-Shot Fine-Grained Image Classification via Multi-frequency Neighborhood and Double-Cross Modulation [EB/OL]. (2022-07-18)[2023-09-05]. https://arxiv.org/abs/2207.08547.

[8] 王永雄，张晓兵. 聚焦-识别网络架构的细粒度图像分类 [J]. 中国图象图形学报，2019，24(4)：23-34. (WANG Y X，ZHANG X B. Fine-Grained Image Classification with Network Architecture of Focus and Recognition [J]. Chinese Journal of Image and Graphics，2019，24(4)：23-34.)

[9] FARRELL R，OZA O，ZHANG N，et al. Birdlets：Subordinate Categorization Using Volumetric Primitives and Pose-Normalized Appearance [C]//Proceedings of the 2011 International Conference on Computer Visio. Piscataway，NJ：IEEE，2011：161-168.

[10] 朱晨鹏，彭宏京，刘学军. 基于双注意力核化双线性网络的细粒度图像分类 [J]. 计算机工程与设计，2022，43(7)：2007-2014. (ZHU C P，PENG H J，LIU X J. Fine Grained Image Classification Based on Dual Attention Kernel Bilinear Networks [J]. Computer Engineering and Design，2022，43(7)：2007-2014.)

[11] 蔡茂，刘芳. 基于细粒度图像分类算法的新冠 CT 图像分类 [J]. 吉林大学学报(信息科学版)，2023，41(4)：676-684. (CAI M，LIU F. CT Image Classification of COVID-19 Based on Fine-Grained Image Classification Algorithm [J]. Journal of Jilin University (Information Science Edition)，2023，41(4)：676-684.)

[12] ZHENG Q H，YANG M Q，ZHANG Q R，et al. A Bilinear Multi-scale Convolutional Neural Network for Fine-Grained Object Classification [EB/OL]. (2018-03-12)[2023-09-05]. https://api.semanticscholar.org/CorpusID:211538378.

[13] 王越，冯振. 基于 CAM 与双线性网络鸟类图像识别方法 [J]. 重庆理工大学学报(自然科学版)，2021，35(11)：136-141. (WANG Y，FENG Z. Bird Image Recognition Method Based on CAM and Bilinear Network [J]. Journal of Chongqing University of Technology (Natural Science Edition)，2021，35(11)：136-141.)

[14] XU K，BA J，KIROS R，et al. Show，Attend and Tell：Neural Image Caption Generation with Visual Attention [EB/OL]. (2015-02-10)[2023-09-06]. https://doi.org/10.48550/arXiv.1502.03044.

[15] 朱丽，王新鹏，付海涛，等. 基于注意力机制的细粒度图像分类 [J]. 吉林大学学报(理学版)，2023，61(2)：371-376. (ZHU L，WANG X P，FU H T，et al. Fine-Grained Image Classification Based on Attention Mechanism [J]. Journal of Jilin University (Science Edition)，2023，61(2)：371-376.)

[16] LIU X X, WANG T, JIANG Y, et al. Fully Convolutional Attention Networks for Fine-Grained Recognition [EB/OL]. (2016-03-22)[2023-09-06]. https://arxiv.org/abs/1603.06765.

[17] ZHONG X, GONG O B, HUANG W X, et al. Squeeze-and-Excitation Wide Residual Networks in Image Classification [C]//2019 IEEE International Conference on Image Processing(ICIP). Piscataway, NJ: IEEE, 2019: 395-399.

[18] YANG X Y, YU C, GAO J X, et al. SAVE: Spatial-Attention Visual Exploration [C]//2022 IEEE International Conference on Image Processing(ICIP). Piscataway, NJ: IEEE, 2022: 1356-1360.

[19] WANG Y M, MORARIU V I, DAVIS L S. Learning a Discriminative Filter Bank within a CNN for Fine-Grained Recognition [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ: IEEE, 2018: 4148-4157.

[20] ZHANG T, CHANG D L, MAI Z Y, et al. Progressive Co-attention Network for Fine-Grained Visual Classification [C]//International Conference on Visual Communications and Image Processing (VCIP). Piscataway, NJ: IEEE, 2021: 2664-2677.

[21] FU J L, ZHENG H L, MEI T. Look Closer to See Better: Recurrent Attention Convolutional Neural Network for Fine-Grained Image Recognition [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ: IEEE, 2017: 4438-4446.

[22] 李晓旭, 安文娟, 武继杰, 等. 通道注意力双线性度量网络 [J]. 吉林大学学报(工学版), 2022, 54(2): 524-532. (LI X X, AN W J, WU J J, et al. Channel Attention Bilinear Metric Network [J]. Journal of Jilin University (Engineering Edition), 2022, 54(2): 524-532.)

[23] GOGUL I, KUMAR V S. Flower Species Recognition System Using Convolution Neural Networks and Transfer Learning [C]//2017 Fourth International Conference on Signal Processing, Communication and Networking (ICSCN). Piscataway, NJ: IEEE, 2017: 1-6.

(责任编辑: 韩 啸)