

衰减解耦机制的图像分类网络

袁 姮¹, 杨继真^{1*}, 张晟翀²

(1. 辽宁工程技术大学软件学院, 辽宁葫芦岛 125105; 2. 光电信息控制和安全技术重点实验室, 天津 300308)

摘要: 针对图像分类网络提取到的特征耦合度高、判别性不足导致关键特征表达能力受限的问题, 本文提出一种衰减解耦机制的图像分类网络(image classification Network with Attenuation Disentangling Mechanism, ADMNet)。首先, 基于生物神经元对信号强度的差异化响应特性, 提出衰减解耦机制, 设计空间衰减解耦(Spatial Attenuation Disentangling, SAD)模块, 将特征图分解至独立子空间并进行不同阈值的衰减变换, 有效解耦提纯关键特征, 过滤空间冗余信息; 设计通道解耦(Channel Disentangling, CD)模块, 使用多尺度一维卷积建模不同范围的通道耦合关系, 动态增强与类别相关的特征通道, 抑制无关通道; 再整合SAD与CD模块, 构建特征衰减解耦(Feature Attenuation Disentangling, FAD)模块, 通过双分支联合操作实现特征的有效解耦, 增强关键特征的判别性, 进而提高图像特征的非线性表达能力。然后, 构建特征聚合池化(Feature Aggregation Pooling, FAP)模块, 通过聚合不同卷积提取到的多尺度特征, 丰富特征表示, 提高空间信息利用率, 减小特征图尺寸。最后, 将FAD模块和FAP模块分别嵌入到残差块的主干路径和残差分支中, 让网络学习耦合度更低、判别性更强的特征以及丰富的上下文信息, 减少网络传递过程中的信息丢失, 提升网络泛化能力。本文方法在CIFAR-10、CIFAR-100、SVHN、Imagenette和Imagewoof数据集上分别取得了96.6%、80.6%、97.5%、89.6%和83.1%的分类准确率。实验结果表明, ADMNet能够有效解耦图像特征, 增强特征的判别性, 减少信息丢失风险, 提升图像分类能力。

关键词: 图像分类; 衰减解耦; 特征解耦; 多尺度特征; 残差分支

基金项目: 国家自然科学基金(No.61601213); 辽宁省教育厅重点基金(No.LJYL049); 辽宁省教育厅重点项目(No.LJ2017QL034)

中图分类号: TP391 **文献标识码:** A **文章编号:** 0372-2112(2026)02-0818-19

电子学报 URL: <http://www.ejournal.org.cn> **DOI:** 10.12263/DZXB.20250940

Image Classification Network with Attenuation Disentangling Mechanism

YUAN Heng¹, YANG Jizhen^{1*}, ZHANG Shengchong²

(1. School of Software, Liaoning Technical University, Huludao, Liaoning 125105, China;

2. Key Laboratory of Optoelectronic Information Control and Security Technology, Tianjin 300308, China)

Abstract: To address the issue of high entangling and insufficient discriminative ability of the features extracted by image classification networks, which limits the expressive power of key features, this paper proposes an image classification network with attenuation disentangling mechanism (ADMNet). Firstly, based on the differential response characteristics of biological neurons to signal intensity, an attenuation disentangling mechanism is proposed: the spatial attenuation disentangling (SAD) module is designed to decompose the feature map into independent subspaces and perform attenuation transformation with different thresholds, effectively disentangling and purifying key features and filtering redundant information; the channel disentangling (CD) module is designed to use multi-scale one-dimensional convolution to model the channel entangling relationships within different ranges, dynamically enhancing feature channels related to classifications and suppressing irrelevant channels; then, by integrating the SAD and CD modules, the feature attenuation disentangling (FAD) module is formed, which achieves effective disentangling of image features through the joint operation of dual-branches, enhancing the discriminability of key features, and thereby improving the nonlinear expression ability of image features. Then, a feature aggregation pooling (FAP) module is constructed. It aggregates multi-scale features extracted by different convolutions, enriches feature representations, improves spatial information utilization, and reduces the size of feature maps. Finally, the FAD and FAP modules are embedded into the main path and residual branches of the residual block, respectively, allowing the network to learn features with lower coupling and stronger discriminability, as well as rich contextual information. This reduces information loss during network transmission and enhances the network's generalization ability. The

method proposed in this paper achieves classification accuracies of 96.6%, 80.6%, 97.5%, 89.6%, and 83.1% on CIFAR-10, CIFAR-100, SVHN, Imagenette, and Imagewoof datasets, respectively. Experimental results show that ADMNet can effectively decouple image features, enhance feature discriminability, reduce the risk of information loss, and improve image classification ability.

Keywords: image classification; attenuation disentangling; feature disentangling; multi-scale features; residual branch

Foundation Item(s): National Natural Science Foundation of China (No.61601213); Project Key Fund of the Liaoning Provincial Department of Education (No.LJYL049); Key Project of the Liaoning Provincial Department of Education (No.LJ2017QL034)

0 引言

图像分类作为计算机视觉领域的核心任务之一,目的是将图像自动归类到预定义类别中,其应用涉及安防监控、医疗诊断、自动驾驶、工业检测以及内容推荐等诸多领域。

最早的图像分类方法通过 SIFT (Scale-Invariant Feature Transform)、HOG (Histogram of Oriented Gradient) 等手工设计算法提取图像局部特征,再使用 SVM (Support Vector Machine)、KNN (K-Nearest Neighbor) 等分类器进行图像分类,这些方法虽然取得了较好的分类效果,但在复杂的图像场景以及大规模数据集上的作用效率低下且效果不佳。为满足自动特征提取以及图像分类性能提升的需要,基于卷积神经网络 (Convolutional Neural Networks, CNN) 的图像分类方法迅速发展。Lecun 等人^[1]提出 LeNet-5,通过卷积层与池化层的交替堆叠实现手写数字识别,有效提取输入图像的平移不变特性。Krizhevsky 等人^[2]提出 AlexNet,采用多层卷积和 ReLU (Rectified Linear Unit) 非线性激活函数自动提取图像的层次化特征,提高特征提取的效率与质量。Simonyan 等人^[3]提出 VGG-Net,通过使用更小的 3×3 卷积核与池化层构建深度卷积神经网络,在保持感受野不变的情况下减少参数数量,验证网络深度与性能之间的正相关关系。Szegedy 等人^[4]提出 GoogLeNet,通过设计多尺度卷积并行计算的 Inception 模块,高效捕捉图像中不同尺度的特征。

随着上述神经网络深度和宽度的不断增加,反向传播过程中的梯度值逐层指数级地减小或增大,梯度消失和梯度爆炸等问题接连涌现。为解决该问题,He 等人^[5]提出 ResNet,通过引入残差连接使信息可在网络层间跳跃传递,实现了前向传播中的恒等映射与反向传播中的无阻碍梯度流动,有效缓解了深层网络的训练难题,提升神经网络的分类性能。

此外,Transformer 的引入催生了一系列新颖的网络架构。Khan 等人^[6]提出 ViT (Vision Transformer),首次成功地将纯 Transformer 架构应用于图像分类任

务,采用基于自注意力机制的全局建模方式,将图像转换为序列化的数据形式,使其能够直接被标准 Transformer 编码器应用,从而产生最终分类效果,具有强大的全局建模能力。Touvron 等人^[7]提出 DeiT (Data-efficient image Transformers),通过引入知识蒸馏策略和改进的训练技术,显著降低 Transformer 对数据量的需求,提高模型的学习效率与泛化能力。Liu 等人^[8]提出 Swin Transformer,通过引入层次化架构与滑动窗口注意力机制,解决 ViT 在处理高分辨率图像时的计算复杂度问题,增强了模型对多尺度特征的捕捉能力。

上述由卷积神经网络和基于 Transformer 的模型主导的图像分类方法,虽然通过一系列的前沿技术(如大规模预训练、自监督学习、对比学习等)提升了图像分类网络的性能、效率和泛化能力,但都忽视了网络提取到的特征图本身携带有大量的冗余信息,而这些信息与类别相关信息纠缠耦合,会降低特征的表达能力和判别性,阻碍网络对图像关键特征的识别,从而影响分类精度。

为应对复杂网络中特征耦合的挑战,特征解耦 (feature disentangling) 方法被提出,其核心目标是将高度耦合的特征表示分解为相互独立的组件,使网络能够分别学习各类特征表示,从而提取出更具判别性的特征信息。Xu 等人^[9]提出对抗性特征解耦网络 (AFD-Net),采用对抗训练与类内重构策略,通过独立编码器解耦图像中的身份特征,但解耦效果受限于特征重构的质量。Liu 等人^[10]则利用编码器分离图像中的域共享与域特有特征,通过全局三元组解耦与实例相似度解耦模块消除域偏差,但相似度分布计算复杂。唐珺琨等人^[11]提出 TRFS (Task Relevant Feature Separation network) 方法,借助注意力一致性学习与正交约束策略解耦域间混合特征,但其正交约束缺乏动态适应性。Yuan 等人^[12]提出多层特征解耦鲁棒网络 (Multi-Level Feature Disentanglement robust network, MLFD),采用最大分类器差异与最小互信息解耦特征,减少特征间统计依赖性,但面临训练不稳定、计算成本高与互信息估计难的问题。

尽管上述方法在特征解耦方面取得了一定进展,但其基于编码器或复杂约束的策略仍存在模块结构复杂、灵活性不足等问题。生物神经系统高效简洁的信号处理机制为此提供了新的思路,其通过信号衰减机制^[13]实现对不同强度的输入实施差异化调节,低响应信号弱抑制,高响应信号强抑制,确保关键信号的高效传递。受此启发,本文模拟生物神经系统多路径并行处理信号的特性,提出了一种衰减解耦机制,该机制通过空间衰减解耦(Spatial Attenuation Disentangling, SAD)与通道解耦(Channel Disentangling, CD)模块构建特征衰减解耦(Feature Attenuation Disentangling, FAD)模块,通过简单的信号衰减函数与高效的一维卷积差异化处理特征信号,抑制冗余特征,减弱其与关键特征之间的耦合,避免了对抗训练、互信息估计等高成本运算,无需引入独立编码器即可实现有效的特征解耦,与先前解决特征耦合的策略相比更为简洁,并且该模块能够根据输入内容进行自适应动态调整,更具灵活性,能够适应多样化的分类场景。基于上述机制,本文提出衰减解耦机制的图像分类网络(image classification Network with Attenuation Disentangling Mechanism, ADMNet)。该网络通过FAD模块从空间与通道两个维度解耦图像特征,增强关键特征的判别性,进而提高特征的非线性表达能力;再通过特征聚合池化(Feature Aggregation Pooling, FAP)模块丰富特征表示,减少信息损失,从整体上提升图像分类性能。

1 衰减解耦机制

在生物神经系统中,单个神经元通过树突接收来自其他神经元的电信号,并在胞体中进行整合,当整合后的信号强度超过特定阈值时,神经元会被激活并产生动作电位。值得注意的是,动作电位的发放模式会随输入信号强度的差异而动态变化。这一机制确保了神经系统能够聚焦于关键信号,对其产生高效且差异化的响应,同时对低于阈值的冗余信号或过度激活的异常信号进行衰减与过滤。CNN中的卷积运算与非线性激活过程,在一定程度上模拟了上述生物神经元的行为,卷积层的加权求和操作对应胞体的信号整合,而激活函数则模拟了基于阈值的放电行为。

然而,现有CNN因为标准卷积操作的平移不变性,在局部感受野内对图像的所有空间位置与特征通道进行等同处理;传统激活函数通常对图像的所有特征施加统一且模式单一的变换,导致网络未能充分模拟神经元对不同强度信号的差异化响应以及对无关信号的衰减过滤这一核心智能机制。这种无差别的处理方式导致生成的特征图中类别相关与无关特征

高度交织耦合,网络无法有效学习图像中的各类特征表示,从而难以自适应地聚焦于最具判别性的图像区域,形成特征表达瓶颈,限制图像分类精度的提升。

经过上述分析可以得出,实现对不同强度特征信号的差异化衰减处理,是过滤无关信息、解耦各类图像特征的一项核心原则。基于此,本文从图像的空间与通道维度出发,提出衰减解耦机制:在空间维度,通过子空间分解与不同阈值的信号衰减函数解耦图像中的各类特征表示,过滤冗余信息,增强关键特征的非线性表达能力;在通道维度,通过多尺度一维卷积建模不同范围的通道耦合关系,动态增强类别相关的特征通道,抑制无关通道;最后,整合空间与通道解耦,形成双分支的特征衰减解耦模块,扩大图像中的特征差异,降低各类特征的耦合度,提升关键特征的判别性,实现特征解耦。

1.1 空间衰减解耦模块

生物神经系统中,神经元通过设定阈值对强、弱信号实施不同强度的衰减,从而过滤无关信号、传递关键信息。在CNN中,卷积提取到的特征图中包含大量与图像类别无关的冗余信息,其与类别相关特征高度耦合,限制关键特征的表达能力。为此,本文设计了空间衰减解耦(SAD)模块。该模块将输入图像划分成多个子空间,并为每个子空间配置具有不同幅度和阈值的信号衰减函数,对各个子空间内的特征实施差异化衰减变换,从而过滤冗余信息,提纯具有不同响应强度的关键特征,最终降低类别相关特征之间的耦合度,增强特征的非线性表达能力。

SAD模块结构如图1所示。

给定输入特征图 $X \in \mathbf{R}^{C \times H \times W}$,其中 H 、 W 分别表示特征图的高度和宽度, C 为通道数。

首先沿通道维度将特征图划分为 n 个子空间,确保每个子空间能够专注学习不同的特征响应模式,有效防止网络对所有图像区域的无差别化处理,为特征解耦提供有效前提条件。每个子空间被分配到的实际通道数如式(1)所示:

$$\text{channel}[i] = \begin{cases} \text{base_size} + 1, & i \leq r \\ \text{base_size}, & \text{otherwise} \end{cases} \quad (1)$$

其中: $i \in \{1, 2, \dots, n\}$ 表示子空间索引; $\text{base_size} = C/n$ 表示特征图被划分到第 i 个子空间中的基础通道数; $\text{Channel}[i]$ 则表示实际通道数,当通道数 C 无法被 n 整除时,将剩余的 r 个通道依次分配至前 r 个子空间中,确保所有通道都能被合理分配,防止因通道分配不均导致的特征表达不充分。

接着,按照上述通道数将影响图像分类的多种特征因素划分至不同子空间。划分的核心依据是特征响应强度的差异:与类别判定强相关的特征通常在网

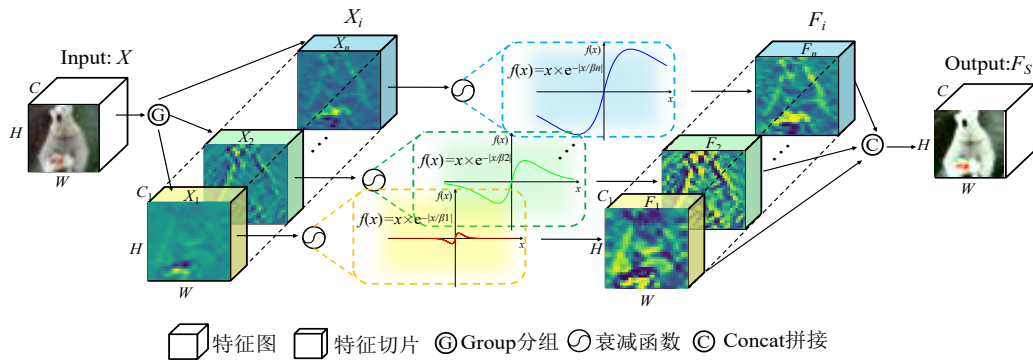


图1 SAD模块结构图

Figure 1 Structure diagram of SAD module

络中会激发更高的响应,表现为中等或较大特征值,例如物体的主体轮廓、边缘等;而与类别无关的冗余信息响应较弱,表现为特征值较小或趋于零,例如图像背景等;此外,还有一些异常激活的极大特征,如噪声等。通过子空间的划分,网络能够实现物理层面的特征解耦,限制图像特征交互范围,降低不同特征信息之间的耦合度,为后续的差异化处理提供结构基础。每个子空间中的特征通道相互关联,视为一个特征切片,所有特征切片的集合构成原始输入特征,如式(2)所示:

$$\{X_i\} = \{X_1, X_2, \dots, X_n\} = X \quad (2)$$

其中: $X_i \in \mathbf{R}^{C_i \times H \times W}$ 表示第 i 个子空间中的特征切片; C_i 表示式(1)得出的第 i 个特征切片被分配的通道数, H 、 W 保持不变。

随后,采用神经元信号衰减函数 $f(x) = x \times \exp(-|x|/\beta)$ 对特征切片实施不同幅度的衰减变换,该函数在不同 β 值下的曲线对比如图2所示。

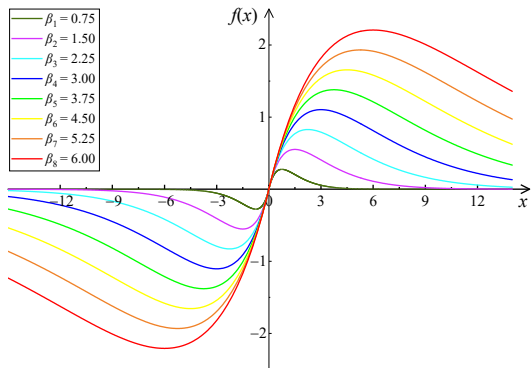


图2 信号衰减函数曲线

Figure 2 Signal attenuation function curves

在该函数中,当特征值 $|x|$ 趋近于0且小于 β 时, $f(x) \approx x$, 衰减函数能够保持小特征值的线性特性,避免过度抑制图像中的潜在特征;当特征值 $|x|$ 逐渐增

大且小于 β 时,衰减函数的非线性效果明显,能够增强图像中关键特征的表达,使类别相关的判别性特征更加突出;当特征值 $|x|$ 大于 β 且持续增大时,衰减函数曲线迅速降低,将特征值衰减为较小值,从而抑制图像中与类别无关的强响应信号。在训练过程中,该函数作为一个可微的非线性变换,其梯度信号会通过反向传播引导网络优化权重。为最小化损失,网络会将高响应特征通道聚集在弱衰减的子空间中,而将低响应或冗余特征通道聚集在强衰减的子空间中。

为实现衰减阈值在不同特征子空间中的差异化分布,本文采用公式化策略设定衰减率超参数 β 。具体而言,将 β 的上限设定为6(见3.4节),并基于此上限,通过公式将 β 值在 n 个子空间上进行线性分配,确保无论 n 取何值,总能产生一组由 $6/n$ 到6均匀递增的衰减参数,使其自适应于子空间的数量变化,如式(3)所示:

$$\beta_i = \frac{6}{n} \times i \quad (3)$$

其中: $i = \{1, 2, 3, \dots, n\}$ 表示子空间索引; $\beta_i = \{\beta_1, \beta_2, \beta_3, \dots, \beta_n\} = \{0.75, 1.5, 2.25, \dots, 6\}$ 表示第 i 个特征切片 $X_i \in \mathbf{R}^{C_i \times H \times W}$ 对应的衰减参数。

随后,对各特征切片施加不同阈值的衰减变换,促使网络在各子空间中独立学习差异化特征,确保不同强度的图像特征都能得到适当表达,降低各类特征间的耦合度。当 $\beta = 0.75, 1.5$ 时,衰减阈值较小,衰减强度大,能够对图像中绝大部分的特征值进行衰减,过滤无关信号的同时仅保留潜在弱特征;当 $\beta = 2.25, 3, 3.75, 4.5$ 时,衰减强度中等,能够过滤部分异常激活信号,并对中等大小的特征值进行不同程度的非线性增强,提高图像中各类特征的判别性;当 $\beta = 5.25, 6$ 时,函数的衰减阈值较大,衰减强度弱,能够对绝大部分特征值实施非线性增强,仅抑制异常强响应的特征信号,尽可能保持原始图像特征完整性,如式(4)所示:

$$F_{c,h,w}^{(i)} = X_{c,h,w}^{(i)} \times \exp\left(-\left|\frac{X_{c,h,w}^{(i)}}{\beta_i}\right|\right) \quad (4)$$

其中: $c \in [1, C]$, $h \in [1, H]$ 和 $w \in [1, W]$ 分别表示通道、高度和宽度索引; $X_{c,h,w}^{(i)}$ 表示第 i 个特征切片的第 c 个通道在第 h 行第 w 列的特征值; $F_{c,h,w}^{(i)}$ 表示对 $X_{c,h,w}^{(i)}$ 进行衰减变换后的输出特征值, 输出特征切片 $F_i \in \mathbf{R}^{C_i \times H \times W}$ 的维度保持不变。

最后, 为维持图像中各类特征表示的相对分布, 将经过衰减变换后的 n 个子特征切片 $F_i \in \mathbf{R}^{C_i \times H \times W}$ 沿通道维度拼接, 确保每个子空间的衰减效果都能在最终特征中得到体现, 整合后的输出 $F_S \in \mathbf{R}^{C \times H \times W}$ 与原输入特征 $X \in \mathbf{R}^{C \times H \times W}$ 维度一致, 如式(5)所示:

$$F_S = \text{Concat}(F_1, F_2, \dots, F_n) \quad (5)$$

综上所述, SAD 将图像特征解耦至独立子空间,

利用不同幅度的信号衰减函数差异化学习各子空间中的特征信息, 对关键特征与冗余特征进行不同程度的衰减解耦, 有效缓解卷积架构中的特征耦合问题, 实现特征提纯并增强特征的非线性表达能力。

1.2 通道解耦模块

除了空间维度上的耦合, 卷积操作为所有特征通道分配相同的重要性, 忽视通道之间的强耦合性以及不同通道在分类任务或图像区域中的贡献差异, 因而无法根据输入内容动态调整通道权重, 实现通道维度的特征解耦。为解决上述问题, 本文设计了通道解耦 (CD) 模块, 使用多尺度一维卷积对不同范围的特征通道间耦合关系进行建模, 使网络聚焦于与类别相关的关键特征通道并为其分配高权重, 抑制无关通道并为其分配低权重, 从通道角度进一步降低特征耦合度。

CD 模块结构如图 3 所示。

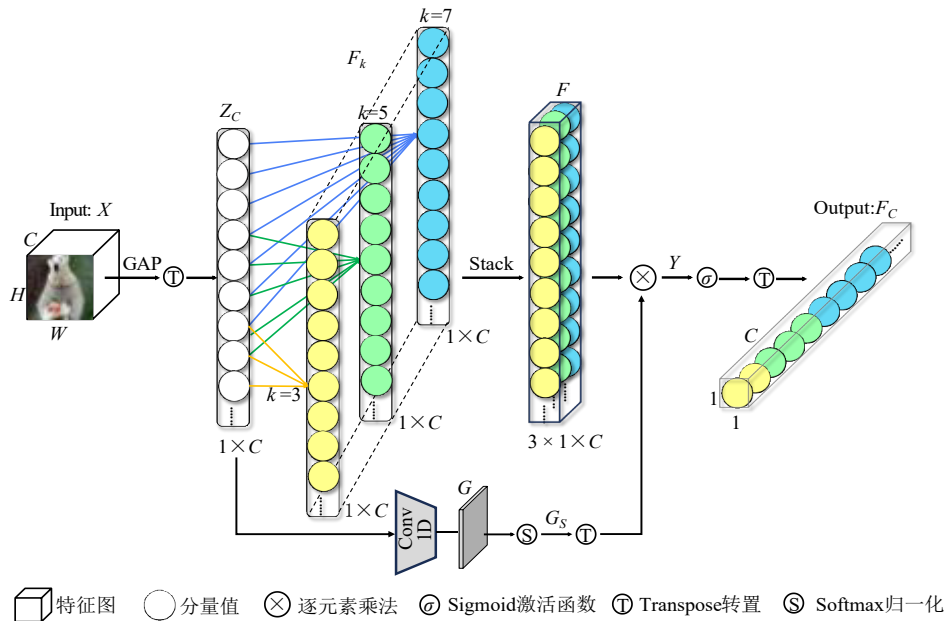


图3 CD模块结构图

Figure 3 Structure diagram of CD module

首先对输入特征 $X \in \mathbf{R}^{C \times H \times W}$ 使用全局平均池化操作, 将特征图的高度 H 与宽度 W 压缩为 1, 用于保留特征图通道维度的全局响应, 去除空间维度干扰并聚焦通道间耦合关系, 为后续进行通道重要性评估提供基础数据, 得到结果 $Z_C \in \mathbf{R}^{C \times 1 \times 1}$, 其计算过程如式(6)所示:

$$Z_C = \text{GAP}(X) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W X(i, j) \quad (6)$$

其中: $\text{GAP}(\cdot)$ 表示通道方向的全局平均池化操作; H 、 W 分别表示输入特征图的高度和宽度; $X(i, j)$ 表示特征图第 i 行第 j 列的特征值。

将 $Z_C \in \mathbf{R}^{C \times 1 \times 1}$ 的维度调整为 $Z_C \in \mathbf{R}^{1 \times C}$, 用于适配后续的一维卷积运算。

随后, 为适应不同层级的图像特征耦合度, 采用多尺度一维卷积对 $Z_C \in \mathbf{R}^{1 \times C}$ 进行建模, 学习其不同局部范围的通道间耦合关系。本文设置三条分支, 其卷积核大小分别为 $K = \{3, 5, 7\}$ 。小尺寸卷积核建模短程通道间的耦合关系, 解耦高度相关的局部特征; 大尺寸卷积核建模长程通道间的耦合关系, 解耦全局特征。所有分支共享同一输入 (即全局平均池化结果), 该过程可表示为

$$F_k = \text{Conv1D}_k(Z_C) \quad (7)$$

其中: $k \in K$ 表示卷积核尺寸;Conv1D(\cdot)表示一维卷积; $F_k \in \mathbf{R}^{1 \times C}$ 表示使用大小为 k 的一维卷积提取到的特征图的通道权重系数。

接着通过增加新维度的方式进行张量拼接,将各分支输出的二维张量 $F_k \in \mathbf{R}^{1 \times C}$ 拼接为一个三维张量 $F \in \mathbf{R}^{3 \times 1 \times C}$,保留各尺度分析结果的独立性,为后续自适应组合提供充分的选择空间,如式(8)所示:

$$F = \text{Stack}(\{F_3, F_5, F_7\}) \quad (8)$$

其中, $F \in \mathbf{R}^{3 \times 1 \times C}$ 表示存储了三条分支权重系数的结果。

同时设计一条简易门控分支,通过一维卷积自适应学习各分支组合的比例系数,较小的比例系数表示在后续进行分支组合时,对该分支提取到的权重系数使用较少。保证网络在浅层特征解耦过程中更依赖小卷积核分支,深层特征解耦过程中更依赖大卷积核分支,该过程可表示为

$$G = \text{Conv1D}_3(Z_C) \quad (9)$$

其中, $G \in \mathbf{R}^{3 \times C}$ 表示门控分支学习到的不同卷积分支自适应组合的比例系数。

对系数 $G \in \mathbf{R}^{3 \times C}$ 沿通道维度执行Softmax归一化,使不同分支相互竞争,促使各分支专注于不同范围的通道耦合关系,减少信息冗余,优化多尺度信息的组合效果,计算过程如式(10)所示:

$$G_S = \text{Softmax}(G, \text{dim} = 1) = \frac{\exp(G_i)}{\sum_j \exp(G_j)} \quad (10)$$

其中,Softmax(\cdot)将系数 $G \in \mathbf{R}^{3 \times C}$ 归一化为一个概率分布向量 $G_S \in \mathbf{R}^{3 \times C}$ (满足 $\sum G_S = 1$)。

将 $G_S \in \mathbf{R}^{3 \times C}$ 维度调整为 $G_S \in \mathbf{R}^{3 \times 1 \times C}$ 后,与存储各分支权重结果的 $F \in \mathbf{R}^{3 \times 1 \times C}$ 进行加权融合,实现对不同大小卷积核提取到的通道权重动态选择与组合,计算过程如式(11)所示:

$$Y = \sum_{k=1}^3 F^k \otimes G_S^k \quad (11)$$

其中: \otimes 表示逐元素相乘; F^k 与 G_S^k 分别表示 $F \in \mathbf{R}^{3 \times 1 \times C}$ 与 $G_S \in \mathbf{R}^{3 \times 1 \times C}$ 第 k 个通道中的系数值; $Y \in \mathbf{R}^{1 \times C}$ 表示加权组合后得到的各特征通道的权重系数。

使用Sigmoid函数将特征图通道权重系数限制在 $[0, 1]$ 范围内。权重值越大,表示该通道与图像的类别相关性越强,网络应该更聚焦于此;权重值越小,则表示该通道冗余或与图像类别判断无关,可以忽略。计算过程如式(12)所示:

$$F_C = \delta(Y) = \frac{1}{1 + \exp(-Y)} \quad (12)$$

其中: δ 表示Sigmoid非线性激活函数; $F_C \in \mathbf{R}^{1 \times C}$ 表示

得到的通道权重。

最后,将 $F_C \in \mathbf{R}^{1 \times C}$ 的维度调整为 $F_C \in \mathbf{R}^{C \times 1 \times 1}$,以便后续与特征图相乘。

综上所述,通过对图像通道间的耦合关系进行建模,能够有效解耦关键特征通道与无关通道,提高特征判别性。

1.3 特征衰减解耦模块

图像的特征表达是一个多维度的综合体现,空间上的关键特征需要由重要的特征通道来传递,而重要的通道也必须在空间上承载有判别性的信息。为构建一个更完备的图像解耦机制,本文综合运用SAD与CD模块,提出了特征衰减解耦(FAD)模块,将空间衰减解耦与通道解耦整合成双分支联合操作的完整解耦结构,如图4所示。

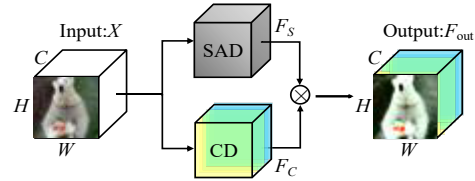


图4 FAD模块结构图

Figure 4 Structure diagram of FAD module

在第一分支中,使用SAD模块对输入特征 $X \in \mathbf{R}^{C \times H \times W}$ 进行空间解耦,衰减过滤冗余信息,从而提纯关键特征,降低它们之间的空间耦合度,得到与原始维度一致的输出特征 $F_S \in \mathbf{R}^{C \times H \times W}$,即

$$F_S = \text{SAD}(X) \quad (13)$$

在第二分支中,使用CD模块对输入特征 $X \in \mathbf{R}^{C \times H \times W}$ 进行通道解耦,扩大特征在通道维度上的差异,得到特征图的通道权重系数 $F_C \in \mathbf{R}^{C \times 1 \times 1}$,即

$$F_C = \text{CD}(X) \quad (14)$$

将空间衰减解耦与通道解耦相结合,有效量化关键特征的通道重要性,增大特征通道间的判别性差异,突出关键特征并抑制冗余特征,从而实现特征的全方位解耦,得到最终输出特征 F_{out} :

$$F_{\text{out}} = F_S \otimes F_C \quad (15)$$

其中, $F_{\text{out}} \in \mathbf{R}^{C \times H \times W}$; \otimes 表示逐元素乘法,将CD分支得到的通道权重 $F_C \in \mathbf{R}^{C \times 1 \times 1}$ 逐元素加权到SAD分支输出特征图 $F_S \in \mathbf{R}^{C \times H \times W}$ 的每个通道上,为类别相关的特征通道赋予高权重,为类别无关的冗余通道赋予低权重。

以类别为“bear”的图像为例,将FAD模块处理前后的特征图进行对比,如图5所示,该图采用伪彩色映射方法表征图像中的特征响应强度,绿色对应弱响应的无关特征,黄色对应强响应的关键特征,其中图5(a)和图5(b)的第一列分别表示原始图像及经FAD模块处

理后的整体效果。在原始特征图[图 5(a)]中,熊的整体轮廓与背景区域的特征响应强度差异较小,均呈现为大面积的绿色或浅黄色,表明特征耦合度高,关键信息不突出。而经 FAD 模块处理后[图 5(b)],熊的整个躯干、头部及四肢区域的特征响应由弱绿色显著增强为亮黄色,与背景区域的差异大幅提升;熊躯干内部的毛发、纹理、边缘轮廓等细节特征从原本模糊的绿色背景中清晰地分离出来,形成局部高响应的亮黄色区域;此外,图像背景区域的响应强度明显衰减,颜色由黄绿色退化为深绿色。上述特征图的转变表明,FAD 模块能够有效地将图像中的主体关键特征从冗余信息中解耦出来,并在过滤无关背景、保留主体信息的同时实现主体细节特征的分离解耦,从而提

高关键特征的判别性。

综上所述,基于 SAD 与 CD 的 FAD 模块,通过双分支结构形成一个“重要性越高,增强效果越显著”的正反馈衰减解耦机制:空间衰减解耦分支过滤冗余信息,有效降低无关特征与关键特征的空间耦合度,提纯并增强类别相关特征;通道解耦分支则动态评估特征图像的通道重要性,为有助于类别判断的关键特征通道分配更大权重,为无关通道分配更小权重。这种联合操作相较于传统 CNN 进一步扩大特征间的差异,使得网络能够有效过滤类别无关特征,学习相互独立性更高的关键特征,最终生成的不再是高度纠缠的混合特征,而是耦合度更低、判别性更强的图像特征。

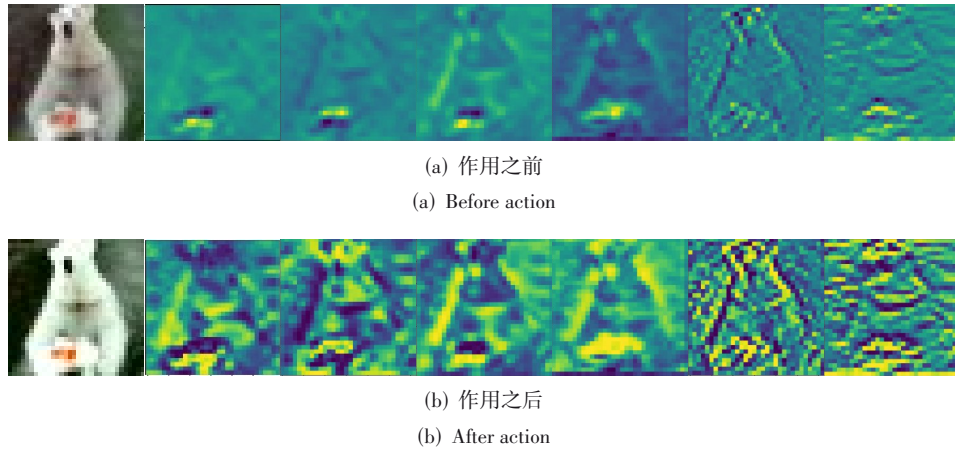


图5 经 FAD 模块作用前后的特征图

Figure 5 The feature maps before and after the action of the FAD module

2 衰减解耦机制的图像分类网络

2.1 特征聚合池化模块

为适应卷积神经网络下采样过程中的特征图尺寸变化,并提升该过程中的空间信息利用率,本文受 ASPP^[14](Atrous Spatial Pyramid Pooling)结构启发,提出特征聚合池化(FAP)模块。该模块在关注图像局部空间特征的同时,融合细节与全局上下文,丰富特征表示,使用平均池化操作聚合多分支结构提取到的特征,从而缩减特征图空间尺寸,最大限度保留多尺度空间特征信息,提高空间信息利用率,缓解因尺度单一而导致的分类不确定性。

FAP 模块结构如图 6 所示,由多分支特征提取路径与平均池化层构成。多分支包括:提供原始像素信息的 1×1 卷积、扩大感受野并增强关键特征响应的门控膨胀卷积,以及协调全局响应的特征重构分支。

首先,在多分支特征提取过程中,FAP 模块对输

入特征 $X \in \mathbf{R}^{C \times H \times W}$ 执行 1×1 卷积操作,提供原始图像的单点特征,保留有效细节信息,得到维度一致的输出特征 $F_{x1} \in \mathbf{R}^{C \times H \times W}$,即

$$F_{x1} = \text{Conv}_{1 \times 1}(X) \quad (16)$$

其中, $\text{Conv}_{k \times k}$ 表示卷积核大小为 $k \times k$ 的二维卷积操作。

门控膨胀卷积分支,通过引入膨胀率(Dilation Rate, r)这一参数,有效扩大卷积核感受野,在特征提取的过程中增强关键特征并抑制无效激活,从而强化局部特征表达,得到维度一致的输出特征 $F_D \in \mathbf{R}^{C \times H \times W}$,即

$$F_D = \text{GDConv}_{r \times r}(X) \quad (17)$$

其中, $\text{GDConv}_{r \times r}$ 表示膨胀率为 r 的门控膨胀卷积操作,其计算过程如式(18)所示:

$$\text{GDConv} = \lambda(\text{Conv}(X)) \odot \delta(\text{Conv}(X)) \quad (18)$$

其中, λ 和 δ 分别表示 ReLU 和 Sigmoid 激活函数,该卷积

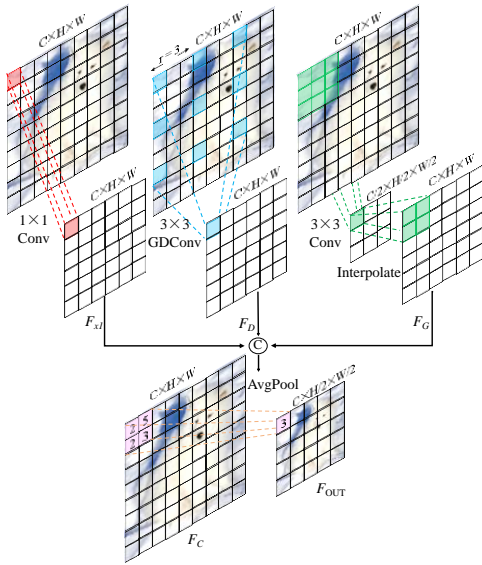


图6 FAP模块结构图

Figure 6 Structure diagram of FAP module

操作获取的特征图实际感受野大小为 $RF = (r \times (k - 1) + 1)^2$ 。

在特征重构分支中,对输入特征图 $X \in \mathbf{R}^{C \times H \times W}$ 依次实施 3×3 卷积进行特征压缩(将特征图的通道数 C 、高度 H 与宽度 W 压缩为原来的 $1/2$)、 1×1 卷积进行通道维度匹配(将通道数 $C/2$ 恢复为 C)、最邻近插值上采样进行空间尺寸的扩大(将高度 $H/2$ 与宽度 $W/2$ 恢复为 H 与 W)、从而将卷积提取到的关键特征响应扩展至全局,得到与原输入维度一致的输出特征 $F_G \in \mathbf{R}^{C \times H \times W}$,即

$$F_G = \infty \left(\text{Conv}_{1 \times 1} \left(\tau \left(\text{Conv}_{3 \times 3} (X) \right) \right) \right) \quad (19)$$

其中, τ 表示 Tanh 双曲正切激活函数,用于对上层 3×3 卷积提取到的特征进行非线性增强,计算公式为

$$\tau = \text{Tanh}(x) = \frac{1 - \exp(-2x)}{1 + \exp(-2x)} \quad (20)$$

该函数的输出范围介于 $(-1, 1)$ 之间,且关于原点对称,零均值特性有助于模型收敛。 ∞ 表示最邻近插值上采样算法,基本原理如图 7 所示。

对于一个待插值的像素点 (X, Y) ,在原始图像中

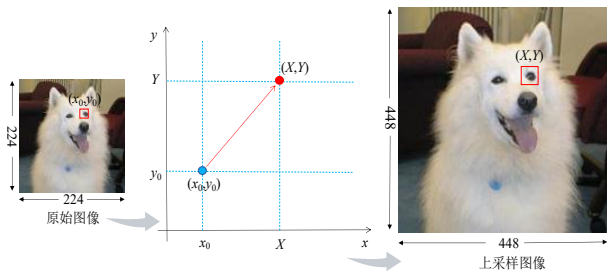


图7 最邻近插值上采样算法

Figure 7 Nearest interpolation upsampling algorithm

取其距离最近的一个已知像素点 (x_0, y_0) 的值作为插值结果。该算法在放大图像时能够较好地保留图像细节,具有计算量小、速度快的特点。

随后,FAP模块将多分支输出特征 $\{F_{x1}, F_D, F_G\}$ 沿通道维度拼接为一个尺寸为 $3C \times H \times W$ 的特征,实现多尺度信息聚合,有效丰富特征表示,再通过 1×1 卷积在不改变特征图空间尺寸 $H \times W$ 的同时将通道数 $3C$ 降为 C ,从而得到结果 $F_C \in \mathbf{R}^{C \times H \times W}$ 。

最后,经平均池化层压缩特征图空间尺寸,在减少参数量的同时保留重要的特征信息,提高空间信息利用率,得到最终输出特征 $F_{OUT} \in \mathbf{R}^{C \times H/2 \times W/2}$,即

$$F_{OUT} = \text{Avg}(F_C) \quad (21)$$

其中, Avg 表示平均池化操作,通过计算特征图中每个 2×2 区域的平均值来降低特征图尺寸。

考虑到多分支输出特征分布差异较大,可能引发神经网络反向传播过程中的梯度异常,本文在平均池化层后加入批量归一化(Batch Normalization, BN)层,将各分支特征分布强制对齐至相近范围,约束梯度幅值。

这种多尺度特征提取聚合后再下采样的操作,能够保留更多有效的多尺度信息,丰富特征表示,提高特征图空间信息利用率,减小特征图尺寸。

2.2 融合FAD与FAP的残差块

在深度卷积神经网络中,特征表示的质量直接影响网络的分类性能。ResNet 残差块(记为 Block,如图 8(a)所示)通过残差分支将原始输入与卷积输出相加,这种设计虽然有效缓解了深层网络中的梯度消失问题,但仍存在两方面不足:

(1)特征耦合度高、判别性受限。固定尺寸的 3×3 卷积对特征图的所有通道与空间位置进行相同处理,引入冗余特征,其与关键特征纠缠耦合,使得网络难以高效学习与类别相关的有效特征。

(2)空间信息丢失。当图像维度发生变化时,ResNet 的下采样残差块(记为 S-Block,如图 8(b)所示)在残差分支中使用步长为 2 的 1×1 卷积进行维度匹配,这一操作使大量与类别判断相关的特征信息丢失。

针对上述 ResNet 残差块的两大局限性,本文将特征衰减解耦(FAD)模块与特征聚合池化(FAP)模块分别嵌入到残差块的主干路径与残差分支中,改进后的结构如图 8(c)D-Block 和图 8(d)P-Block 所示。

具体改进包括以下两方面:

(1)主路径增强。在原有两个 3×3 卷积层后嵌入 FAD 模块,该模块通过多尺度一维卷积动态建模通道间耦合关系,并利用多阈值的信号衰减函数对图像特征进行解耦分离,实现对上层卷积输出的特征进行过滤与提纯,进而使网络学习耦合度更低、判别性

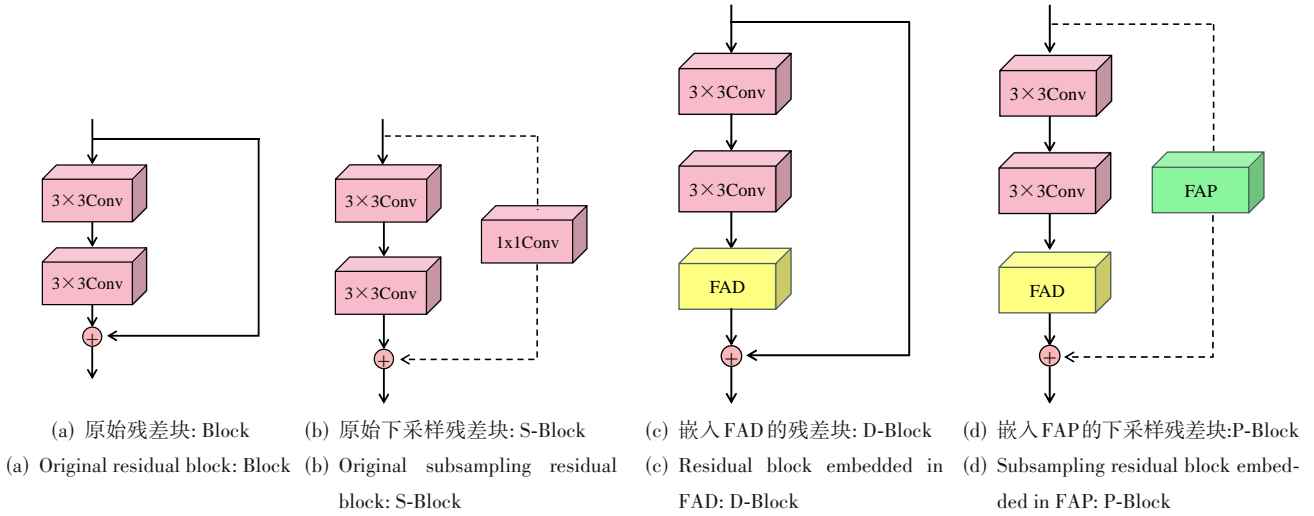


图8 嵌入FAD与FAP的残差块

Figure 8 Residual blocks embedded with FAD and FAP

更强的特征。

(2)残差分支改进。使用FAP模块替代原步长为2的1×1卷积,该模块通过点卷积、门控膨胀卷积和特征重构器对特征图的多尺度空间信息进行有效学习与提取,再经平均池化压缩特征图尺寸,在实现下采样的同时极大程度地保留更加丰富的上下文信息,显著提高特征图空间信息利用率,缓解 ResNet 残差网络下采样过程中的信息丢失问题。

原始 ResNet 的残差映射关系为

$$\text{Res}(x)=F(x)+x \quad (22)$$

其中,Res(x)表示待拟合对象,F(x)表示残差函数,x表

示学习输入。

本文改进后的残差映射关系为

$$N(x)=D(x)+P(x) \quad (23)$$

其中,N(x)表示待拟合对象,D(x)表示经过FAD模块处理后的残差输出,P(x)表示经过FAP模块处理后的输入特征。

2.3 网络总体架构

本文提出的衰减解耦机制的图像分类网络主要由图像预处理模块、浅层特征提取模块、融合FAD与FAP的残差模块以及分类输出模块构成,整体架构如图9所示。

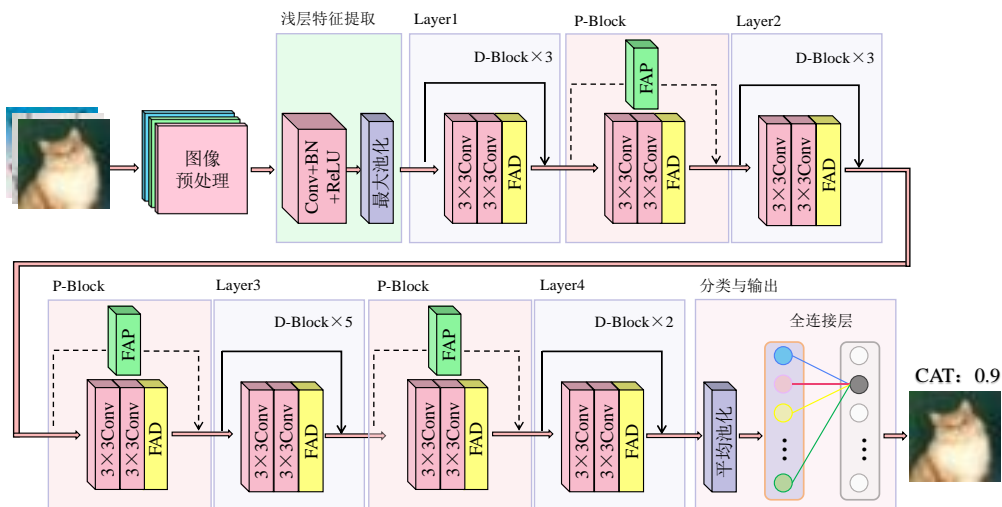


图9 ADMNet总体架构图

Figure 9 Overall architecture diagram of ADMNet

ADMNet的主干网络包含以下四个阶段。

(1)图像预处理。对输入图像执行多类型数据增

强,具体包括:随机水平翻转,提升数据多样性,使模型能够学习镜像不变性特征;边缘填充结合随机裁

剪,在保留图像边缘信息的同时避免目标主体被截断;随机遮挡,通过在图像中随机生成方形掩码区域,遮挡局部区域,减弱模型对局部特征的过度依赖,使其关注全局上下文信息。

(2)浅层特征提取。使用一组标准 7×7 卷积、批量归一化(BN)与激活单元(ReLU)对预处理后的图像进行浅层特征提取,拓展输入图像的特征通道数,增加特征维度。在处理小尺寸数据集时,为保留更多空间信息,修改卷积核尺寸为 3×3 ,并移除最大池化层。

(3)深层特征提取。将上一阶段输出的特征图依次输入至 Layer1、Layer2、Layer3、Layer4 中,提取耦合度更低、判别性更强的深层特征,每个 Layer 由多个融合了 FAD 的残差块 D-Block 构成,每个 Layer 之间通过融合了 FAP 的下采样残差块 P-Block 匹配特征图的维度变化,在丰富上下文信息的同时,减少网络传递过程中的信息损失。

(4)分类输出。通过全局平均池化将特征图压缩为一维向量,用于聚合全局特征信息,随后利用全连接层将特征映射到类别空间(输出维度等于数据集类别数目),并通过 Softmax 函数生成类别概率分布,最终输出图像分类结果。

3 实验与结果分析

3.1 实验环境

本文实验在 Windows11 操作系统上开展,利用 Python 语言和 Pytorch 框架构建网络模型 ADMNet,计算硬件为 24 GB 显存的 NVIDIA RTX 3090 显卡。为评估模型性能,实验选取 CIFAR-10、CIFAR-100、SVHN、Imagenette 以及 Imgewoof 作为数据集,涵盖不同类型的分类任务数据集属性见表 1。

表 1 实验数据集

Table 1 Experimental datasets

数据集	尺寸/像素	类别	训练集	测试集
CIFAR-10	32×32	10	50 000	10 000
CIFAR-100	32×32	100	50 000	10 000
SVHN	32×32	10	73 257	26 032
Imagenette	224×224	10	9 469	3 925
Imgewoof	224×224	10	9 025	3 929

在训练过程中,将总迭代轮次设定为 200,确保模型能够充分学习数据规律并收敛。优化策略选取随机梯度下降(Stochastic Gradient Descent,SGD)作为优化器,并配置如下超参数:动量(Momentum)设置为 0.9,加速模型的收敛过程;权重衰减(Weight Decay)系数设置为 5×10^{-4} ,作为一种有效的正则化手段抑制模型过拟合;学习率调度采用分步衰减策略,初始学习率

(Learning Rate)取 0.1,并设定在训练轮次达到 60、120 和 160 时,依次将学习率乘以衰减系数 0.2,有助于模型在初期快速收敛,并在后期进行精细调优。此外,为进一步提升训练效果,采用交叉熵损失(Cross-Entropy Loss)函数,并融合标签平滑(Label Smoothing)技术(平滑系数为 0.1),用于减轻模型过拟合并增强其泛化能力;网络的所有权重均使用 Kaiming 正态分布进行初始化。针对不同数据集的分辨率与规模,设置差异化的批次大小(Batch Size)以提升训练效能,在分辨率较低的 CIFAR-10、CIFAR-100 及 SVHN 数据集上将批次大小设定为 128,而对于分辨率更高的 Imagenette 与 Imgewoof 数据集,则将批次大小设置为 64。

本文实验基于经过基础数据增强处理的数据集展开,并以验证集的分类准确率(Accuracy, Acc)作为模型性能的核心评价指标。

3.2 FAD 模块嵌入位置及数量实验

特征衰减解耦模块在卷积神经网络中的嵌入位置会显著影响模型性能。不恰当的嵌入位置可能会引发以下问题:(1)引入冗余参数,增加过拟合风险;(2)破坏卷积层的局部归纳偏置,降低特征提取效率。为研究 FAD 模块在 ResNet 残差块中的最优嵌入策略,本文设计了 7 种嵌入方案,各方案所对应的模块结构如图 10 所示。

在统一训练策略以及控制其他变量一致的前提下,使用不同的数据集 CIFAR-10、CIFAR-100、SVHN、Imagenette 与 Imgewoof 在 ResNet-34 网络模型中进行训练与验证,实验结果如图 11 所示。

实验结果表明,Post-Conv 方案在五个数据集上均取得最高分类准确率,分别为 96.6%、80.5%、97.4%、88.4% 与 80.9%。尽管在某些数据集中,插入多个 FAD 模块也可以达到相近准确率,但会导致模型冗余,进而增大过拟合风险。因此,将 FAD 模块嵌入到残差块的第二层卷积之后,通过对上层卷积提取的特征进行衰减解耦与提纯,能够过滤冗余信息并差异化增强图像中关键特征的非线性表达能力。

3.3 SAD 中子空间数量 n 实验

为网络选择合适的超参数可以有效提升分类性能,选择不合适的超参数反而会阻碍网络的性能。在空间衰减解耦(SAD)模块中,子空间数量 n 作为核心超参数,直接影响该机制的特征解耦效果。若子空间数量过少,所有特征通道共享相同衰减变换,难以区分不同层级特征;相反,若子空间数量过多,则空间内特征通道数过少,会破坏特征间的关联性。

为探究 SAD 中子空间数量 n 对网络分类性能的影响,本文在五个数据集上对 n 取值由 2~16 进行实验,结果如图 12 所示。

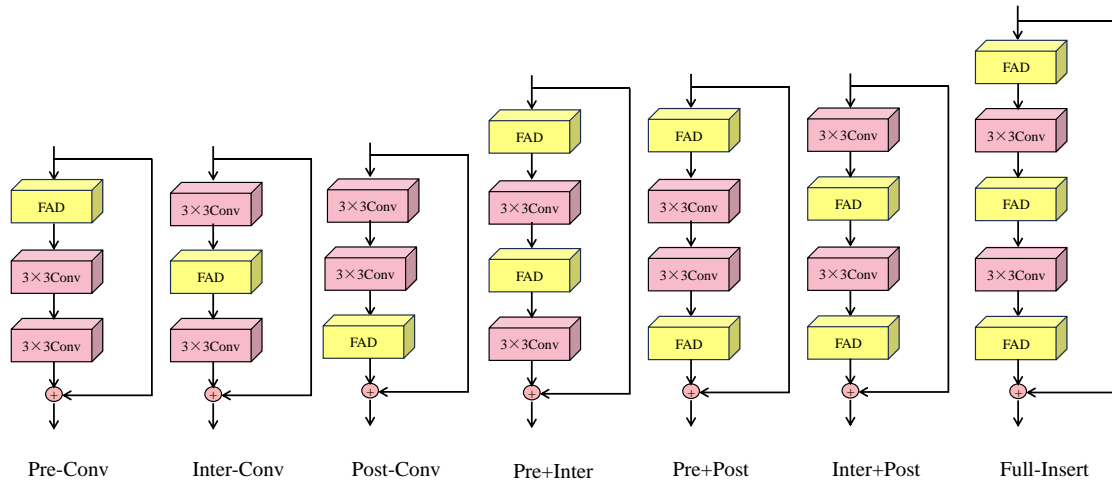


图 10 FAD 模块嵌入残差块的不同方案

Figure 10 Different schemes of the FAD module embedding residual blocks

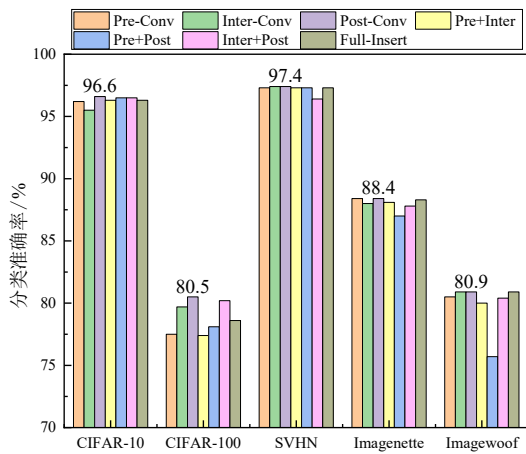
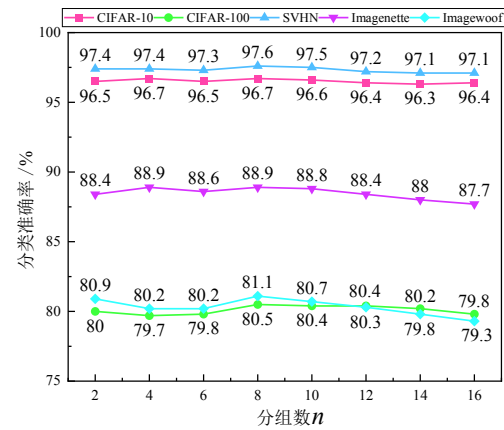


图 11 不同嵌入方案在实验数据集上的准确率柱状图

Figure 11 Accuracy histogram of different embedding schemes on experimental datasets

实验结果表明,图像分类网络的分类准确率并未随子空间数量 n 的增加而持续提升,而是在 $n=8$ 时达到最优,此时网络在数据集 CIFAR-10、CIFAR-100、SVHN、Imagenette 与 Imagewoof 上的分类准确率分别为 96.7%、80.5%、97.6%、88.9% 与 81.1%,随着 n 的进一步增大,网络的分类准确率持续下降或达到饱和状态。

因此,当 n 过小时,所有特征切片共享一个或少数几个 β 值,特征子空间区分度不足,SAD无法差异化处理各类空间特征,导致图像解耦效果受限;当 n 过大时,子空间内部特征稀少,过少的特征通道无法有效映射形成高级特征表示,从而减弱了特征之间的必要关联,影响特征的表达能力。 $n=8$ 的设置能够较好地平衡各个特征切片间的特异性与各个特征切片内的特征交互性,在实现有效的图像特征解耦的同时,保持必要的特征融合,从而显著提升图像分类性能。

图 12 不同分组数 n 对准确率的影响Figure 12 The impact of different numbers of groups n on accuracy

3.4 SAD 中衰减率 β 实验

在 SAD 中,当子空间数量固定时,衰减率参数 β 直接控制衰减变换的处理强度。 β 值越小,衰减强度越剧烈,对激活值的抑制越严格; β 值越大,函数允许通过的特征值越大,对过度激活值的抑制能力越弱。因此,选择合适的 β 值对 SAD 的性能至关重要。为确定最优 β 值,在子空间数量 $n=8$ 的基础上,于五个数据集上对 β 取值上限由 3~9 进行实验,准确率变化如图 13 所示。

实验结果表明,当 $\beta=6$ 时,网络在 CIFAR-10、CIFAR-100、SVHN、Imagenette 与 Imagewoof 数据集上取得最高分类准确率,分别为 96.7%、80.5%、97.5%、88.7%、80.9%。因此,当 β 值过小时,衰减函数对绝大多数激活值均产生强烈抑制,导致解耦过度,有效特征丢失;当 β 值过大时,变换近似于线性操作,模块无法差异化学习各类特征表示,导致解耦不足,限制模型性能的提升。而 $\beta=6$ 的设置抑制冗余与保留有

效信息之间达到最佳平衡,能够对无关背景这类小特征值进行一定程度的抑制,同时保留甚至增强代表边缘、纹理等中等强度的特征值,从而得出更加纯净与更具表达能力的特征,实现特征解耦,提升图像分类性能。

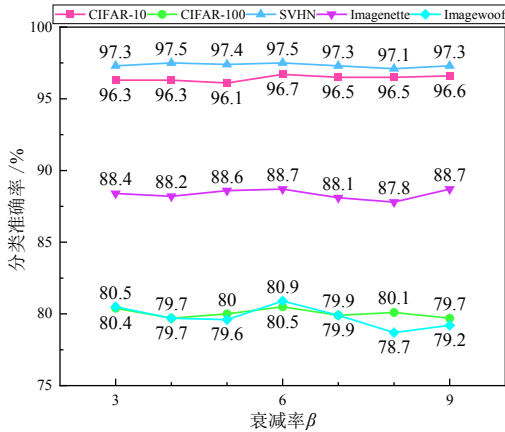


图 13 不同衰减率 β 对准确率的影响

Figure 13 The impact of different β on accuracy

3.5 双分支联合作用对比实验

FAD 模块旨在通过空间衰减解耦(SAD)分支与通道解耦(CD)分支的联合作用,增强网络的特征提取能力。为验证其有效性,本文以 ResNet-34 为基础网络,设计以下 3 组对比实验:组 1(仅保留 SAD 分支,移除 CD 分支)、组 2(仅保留 CD 分支,移除 SAD 分支)、组 3(完整 FAD 模块)。在 CIFAR-10、CIFAR-100、SVHN、Imagenette 和 Imgewoof 数据集上的实验结果如图 14 所示。

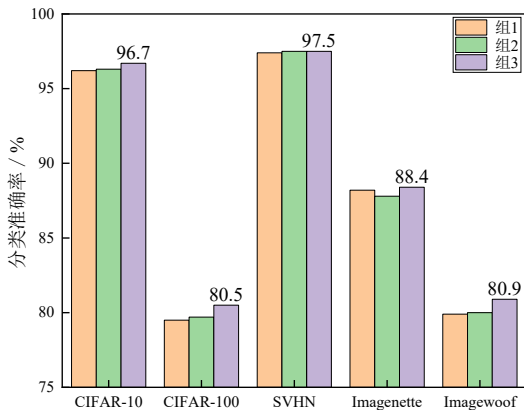


图 14 不同组别在实验数据集上的准确率柱状图

Figure 14 Accuracy histogram of different groups on experimental datasets

实验结果表明,完整 FAD 模块(组 3)在五个数据集上均取得最高分类准确率。具体而言,组 1 仅使用

空间衰减解耦模块,虽然能够有效抑制冗余信息并差异性增强关键特征,缓解全局过度激活导致的特征耦合问题,但缺乏对关键特征的进一步聚焦机制;组 2 仅采用通道解耦模块,赋予高密度信息通道更高权重,但某些因无关细节而产生强响应的特征通道可能被误判为重要通道并增强,导致冗余信息被放大。

FAD 模块中双分支联合作用,有效克服了上述单分支的局限。SAD 分支解决特征质量问题,通过特征衰减解耦与提纯,为通道解耦分支提供高质量、低冗余的特征输入;CD 分支解决特征重要性问题,通过自适应分配权重,引导网络聚焦于最具判别性的特征通道。二者共同构成一个高效的特征衰减解耦模块,确保网络准确识别低耦合度特征并提高关键特征判别性,全面提升图像分类性能。

3.6 FAP 模块嵌入位置与数量实验

ADMNet 主干网络包含 4 个层级(Layer),每层级间均通过下采样残差块(S-Block)实现网络加深过程中特征图维度匹配。传统 ResNet 使用步长为 2 的 1×1 的卷积进行下采样映射,该操作会导致显著的特征信息丢失,进而影响网络后续推理与预测的准确性并降低模型泛化能力。

基于此,本文采用 FAP 模块替换 S-Block 中的 1×1 卷积,构建新的下采样残差块 P-Block。为研究 P-Block 的嵌入位置与数量对网络性能的影响,本文共设计 8 种排列组合方式,如图 15 所示。各组合方式在多个数据集上的分类准确率结果如图 16 所示。

实验结果表明,当将所有 S-Block 替换为基于 FAP 模块的 P-Block(组合方式 H)时,模型在五个数据集上分别达到 96.7%、80.1%、97.5%、89.5% 与 82.7% 的分类准确率,性能均优于其他组合。证明 P-Block 通过特征聚合与多尺度信息保留机制,有效缓解下采样过程中的信息丢失问题,丰富特征表达,显著提升模型在不同规模与复杂度数据集上的泛化能力。因此,采用 P-Block 作为下采样残差块是提升 ADMNet 性能的有效策略,也为后续网络结构优化提供重要的实验依据与设计参考。

3.7 FAP 中不同膨胀率实验

在 FAP 模块的门控膨胀卷积分支中,膨胀率 r 是关键超参数,直接影响模型感受野大小与特征提取能力。较小的膨胀率有助于保持较高的空间分辨率,适合捕捉局部特征;而较大的膨胀率能够扩大感受野覆盖范围,有利于提取全局上下文信息。

为探究膨胀率对网络分类准确率的影响,本文设置了膨胀率取值为 2、3、4、5、6 的五组实验,实验选用 CIFAR-100、CIFAR-10、SVHN、Imagenette 及 Imgewoof 作为训练与测试数据集,并以分类准确率最高值为依

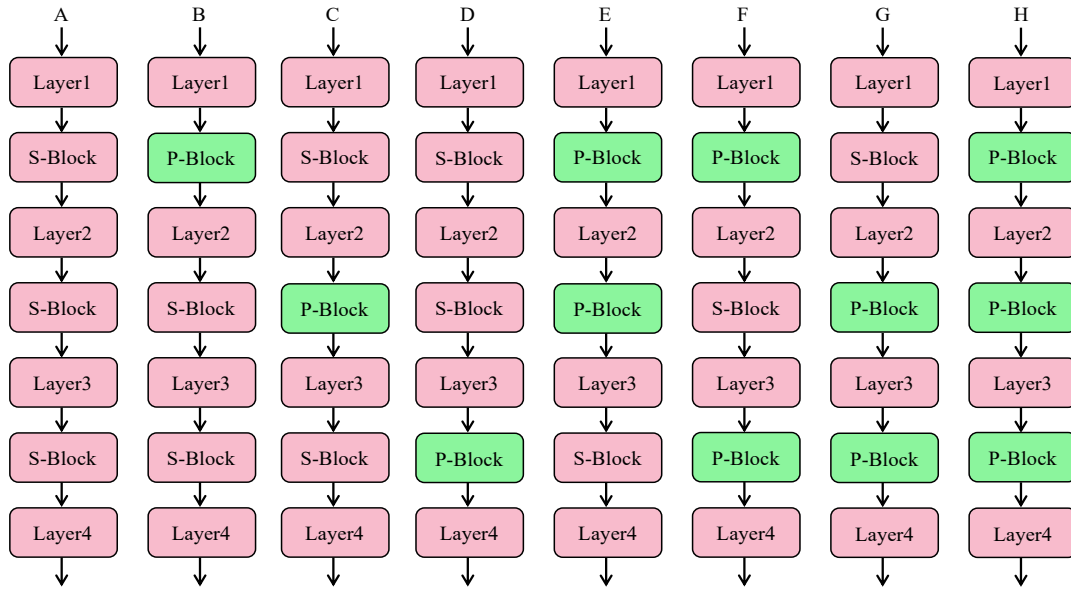


图 15 嵌入 P-Block 的 8 种排列组合方式

Figure 15 8 types of combinations of embedded P-Block

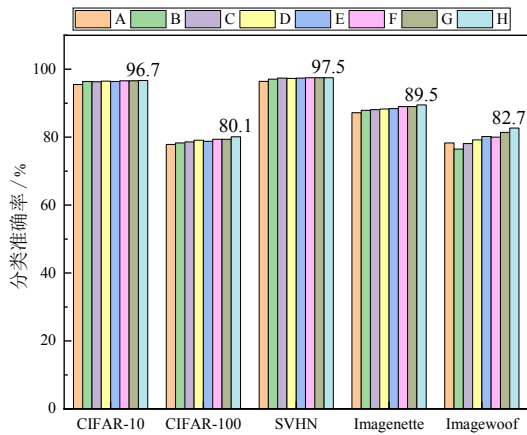


图 16 不同组合方式在实验数据集上的准确率柱状图

Figure 16 Accuracy histogram of different combinations on experimental datasets

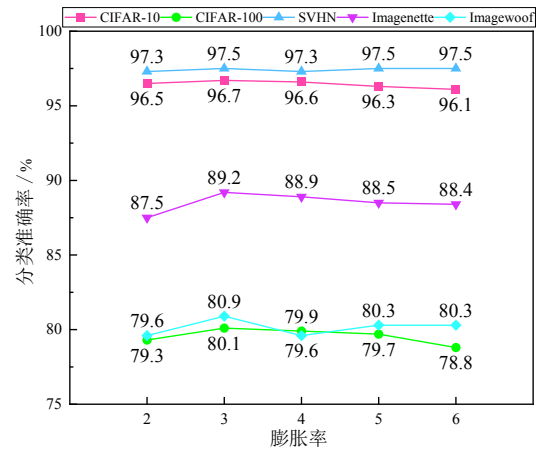


图 17 不同膨胀率对准确率的影响

Figure 17 The impact of different dilation rates on accuracy

据确定最优膨胀率,图 17 展示不同膨胀率下网络在五个数据集上的分类准确率。

实验结果表明,当膨胀率为 3 时,模型在所有数据集上均取得最优分类性能。这是因为过小的膨胀率难以捕获足够的上下文信息,而过大的膨胀率则容易导致特征值之间的相关性丢失,产生网格效应。膨胀率为 3 的设置能够较好地平衡局部特征与全局信息的提取,在避免小膨胀率感受野有限、局部信息过强问题的同时,缓解大膨胀率可能引发的细节丢失问题,从而能够捕获更丰富的局部细节与上下文信息,提高特征利用率。

3.8 消融实验

为验证 ADMNet 中各模块的有效性,本文在选取

的五个数据集上设计并开展了消融实验。实验中定义如下网络:Net1 表示在 ADMNet 中移除 FAP 模块,Net2 表示在 ADMNet 中移除 FAD 模块,Net3 表示修改首层卷积核尺寸为 7×7 并保留最大池化层,BaseNet 表示原始 ResNet-34 架构,作为性能基线。各网络消融实验的具体结果如表 2 所示。

实验结果表明,ADMNet 在五个数据集上均取得了最优性能,准确率分别为 96.6%、80.6%、97.5%、89.6% 和 83.1%,与 BaseNet 相比,ADMNet 在各数据集上的准确率分别提升了 6.9、17.7、2.1、2.4 和 4.8 个百分点,证明 FAD 与 FAP 模块的联合嵌入能够有效提升图像分类网络的性能。

与 Net1 相比,ADMNet 在 CIFAR-10、CIFAR-100、SVHN、Imagenette 和 Imagewoof 数据集上分别提升了

表 2 ADMNet的消融实验结果 单位:%
Table 2 Ablation experiment results of ADMNet unit: %

网络	CIFAR-10	CIFAR-100	SVHN	Imagenette	Imagewoof
ADMNet	96.6	80.6	97.5	89.6	83.1
Net1	95.5	79.7	97.1	88.4	80.9
Net2	95.2	78.6	96.8	88.1	82.7
Net3	96.1	80.1	97.3	—	—
BaseNet	89.7	62.9	95.4	87.2	78.3

注:加粗字体表示各列最优结果。

1.1、0.9、0.4、1.2 和 2.2 个百分点,证明了FAP模块通过多尺度特征聚合与保留,有效减少了采样过程中的信息丢失。与Net2相比,ADMNet在各数据集上分别提升了1.4、2.0、0.7、1.5 和 0.4 个百分点,表明FAD模块通过空间衰减解耦分支与通道解耦分支的联合作

用,能够有效解耦图像中的各类特征表示、抑制冗余信息、增强关键特征判别性。Net3在Imagenette和Imagewoof上未取得有效结果,而在其余小尺寸数据集上性能仍低于ADMNet,说明首层小卷积核与移除最大池化层的设计更适配小尺寸图像的特征提取,有助于保留更多空间细节信息。

各网络在不同数据集上的训练曲线如图 18 所示。无论是在图像尺寸较小或较大,类别数量较少或是较多的数据集上,ADMNet的收敛速度更快且最终准确率更高,进一步验证了FAP与FAD模块对于图像分类任务的有效性:FAP模块减少网络传递过程中的特征信息流失,FAD模块增强类别相关特征的判别性与非线性表达能力,二者共同作用使ADMNet能够更有效地学习耦合度低的判别性特征,提升图像分类能力。

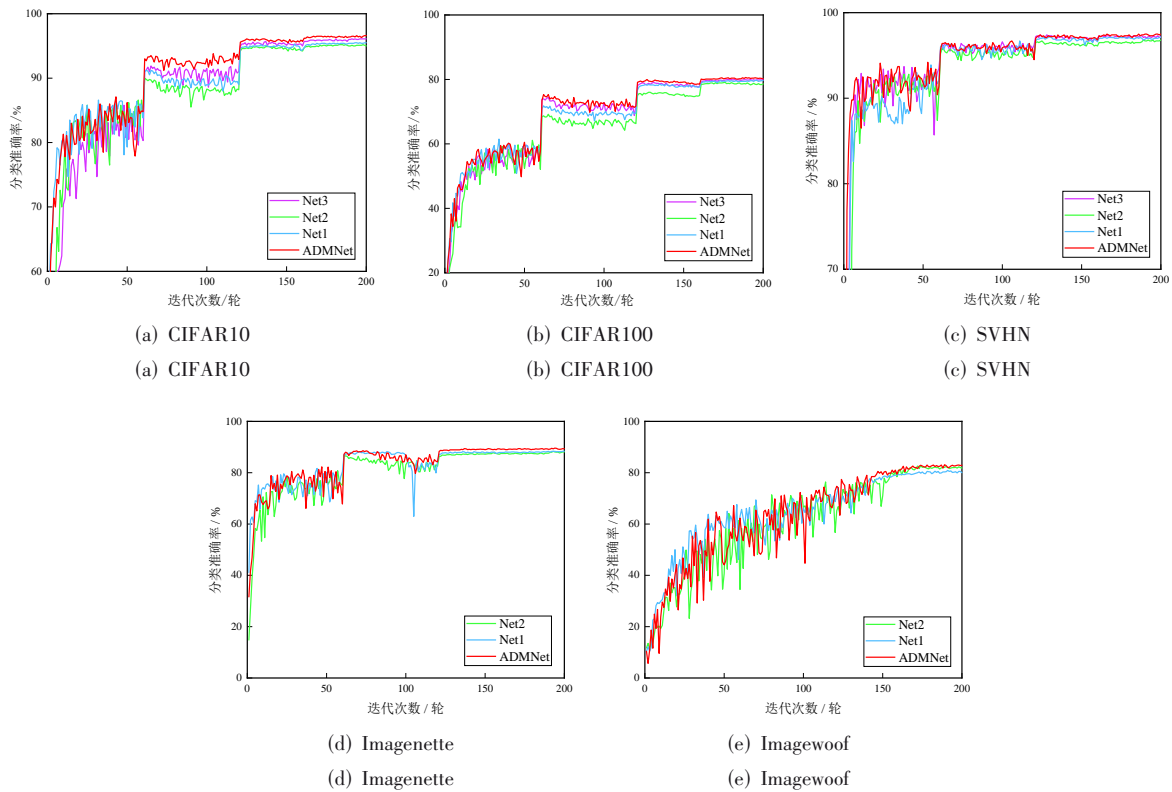


图 18 各个网络在五种数据集上的分类准确率

Figure 18 Classification accuracy of each network on five datasets

3.9 对比实验

为验证ADMNet的先进性与有效性,本文选取多个图像分类模型在五个数据集上进行了广泛的对比实验。

参与对比的模型包括:经典残差网络ResNet-34,高效轻量型网络Efficient-Nets^[15]与GhostNet^[16],高性能网络DenseNet^[17]、QKFormer^[18]与WideResnet-28-10^[19],以及近些年提出的多种先进架构,如Couplformer^[20]、SimpleNetv2^[21]、GAC-SNN^[22]与MMA-CCT-

7/3 × 2^[23],专注于特定任务优化的网络TA-DFKD^[24]、AugLocal^[25]、FAVOR+^[26]、ATONet^[27]和FDPRNet^[28]。

为确保结果的可比性与可靠性,本文所有对比数据均通过以下两种方式获取:

(1)对于未开源代码的网络模型,直接引用原始论文中报告的实验结果。

(2)对于提供了开源代码的网络模型,通过复现实验获得准确率。

各网络在五个数据集上的分类准确率对比如表 3 所示。由表 3 数据可得出,ADMNet 在五个数据集上

均取得最高分类准确率,显著优于其他 15 个对比模型,验证了其在图像分类任务中的有效性。

表 3 各网络在五个数据集上的分类准确率

单位: %

Table 3 Classification accuracy of various networks on five datasets

unit: %

网络	发表年份	CIFAR-10	CIFAR-100	SVHN	Imagenette	Imagewoof
ResNet-34	2015 年	95.54	77.86	96.43	86.75	78.30
EfficientNets	2019 年	94.01	75.96	93.32	88.01	77.93
GhostNet	2020 年	94.92	77.15	93.86	87.83	78.22
DenseNet	2017 年	94.24	78.84	94.95	87.72	77.91
QKFormer	2024 年	96.18	80.26	97.13	88.32	81.65
WideResnet-28-10	2016 年	95.83	79.50	95.21	88.34	78.71
Couplformer	2023 年	93.54	73.92	94.26	87.91	77.89
SimpleNetv2	2018 年	95.97	80.29	96.49	—	—
GAC-SNN	2024 年	96.00	80.23	96.17	—	—
MMA-CCT-7/3×2	2023 年	94.74	77.50	—	—	—
TA-DFKD	2024 年	95.46	78.42	96.83	87.62	80.57
AugLocal	2024 年	95.51	79.23	96.91	87.93	80.89
FAVOR+	2020 年	91.42	72.56	93.21	88.16	77.57
ATONet	2024 年	94.51	78.54	95.21	86.67	80.19
FDPRNet	2025 年	95.70	80.01	96.96	—	—
ADMNet	2025 年	96.60	80.60	97.50	89.60	83.10

注:加粗字体表示各列最优结果。

为进一步定量评估 ADMNet 在图像分类任务中的性能,图 19 对比展示了 ResNet-34 与 ADMNet 在 CIFAR-10(图像尺寸 32×32)和 Imagewoof(图像尺寸 224×22)数据集上的分类混淆矩阵。图 19 标注了网络对数据集中各个图像类别的正确或错误预测的样本数量,主对角线上的数值代表网络正确分类的图像样本数量,数值越大表明网络的分类精度越高;非主对角线数值则代表网络错误分类的混淆样本数量,数值越大表明网络的分类精度越低。

分析可知,ADMNet 的混淆矩阵呈现出两个显著优势:其一,其主对角线上的数值普遍高于 ResNet-34,表明 ADMNet 在所有图像类别上均取得更高的正确分类样本数,即整体分类精度更高;其二,其非主对角线上的错误数值更低,且分布更为分散与均匀,未出现明显的错误集中现象,表明 ADMNet 的分类错误模式得到改善,决策边界更为清晰。

这一结果印证了本文方法的有效性:通过 FAD 模块在空间与通道维度上的双重解耦,有效净化特征表示,增强关键特征的判别性,从而显著提升模型对类间差异的捕捉能力,降低相似类别间的误判率;通过 FAP 模块聚合多尺度特征,提取丰富的上下文信息,有效减少因空间信息丢失或尺度单一而导致的分类不确定性;再通过 FAD 与 FAP 模块的联合作用有效解耦图像特征,提升特征的类间判别力,从而使 ADMNet 具备更出色的图像分类能力。

3.10 热力图可视化分析

在图像分类任务中,热力图通过颜色编码直观呈现模型对图像不同区域的关注程度,红色高权重区域标识关键特征(如物体的典型轮廓),蓝色低权重区域对应贡献较小的背景或冗余信息,能够有效显示模型的决策依据,增强网络的可解释性。

为验证本文网络 ADMNet 在聚焦低耦合、高判别性特征上的有效性,本文将其与五种主流分类网络 ResNet-34、ShuffleNet^[29]、MobileNet_V2^[30]、VGGNet、GoogleNet 进行热力图对比实验。图 20 展示了各网络生成的热力图可视化结果,其中前三张图片选自 CIFAR-10 数据集,后三张图片选自 Imagewoof 数据集。

根据图 20 的可视化结果可以看出,ADMNet 的关注区域更集中、更精确,其空间权重分布更贴近生物视觉系统的特征聚焦机制,能够精准聚焦图像中的类别主体并为其赋予高权重,表现为猫的头、鸟的躯干以及犬的四肢等部位呈现出连贯的亮红色区域;同时,对无关信息显著抑制并赋予其低权重,表现为图像中的天空背景、草地环境等呈现出大面积的深蓝色区域。而其他模型的关注区域较大,其红色与黄色区域覆盖了较多无关的背景信息,尤其是轻量化模型 ShuffleNet 与 MobileNet 热力图中的红色区域呈现弥散分布,未能有效聚焦图像中的主体特征。

为量化分析 ADMNet 生成热力图的准确性与可靠性,引入最相关优先(Most Relevant First, MoRF)扰

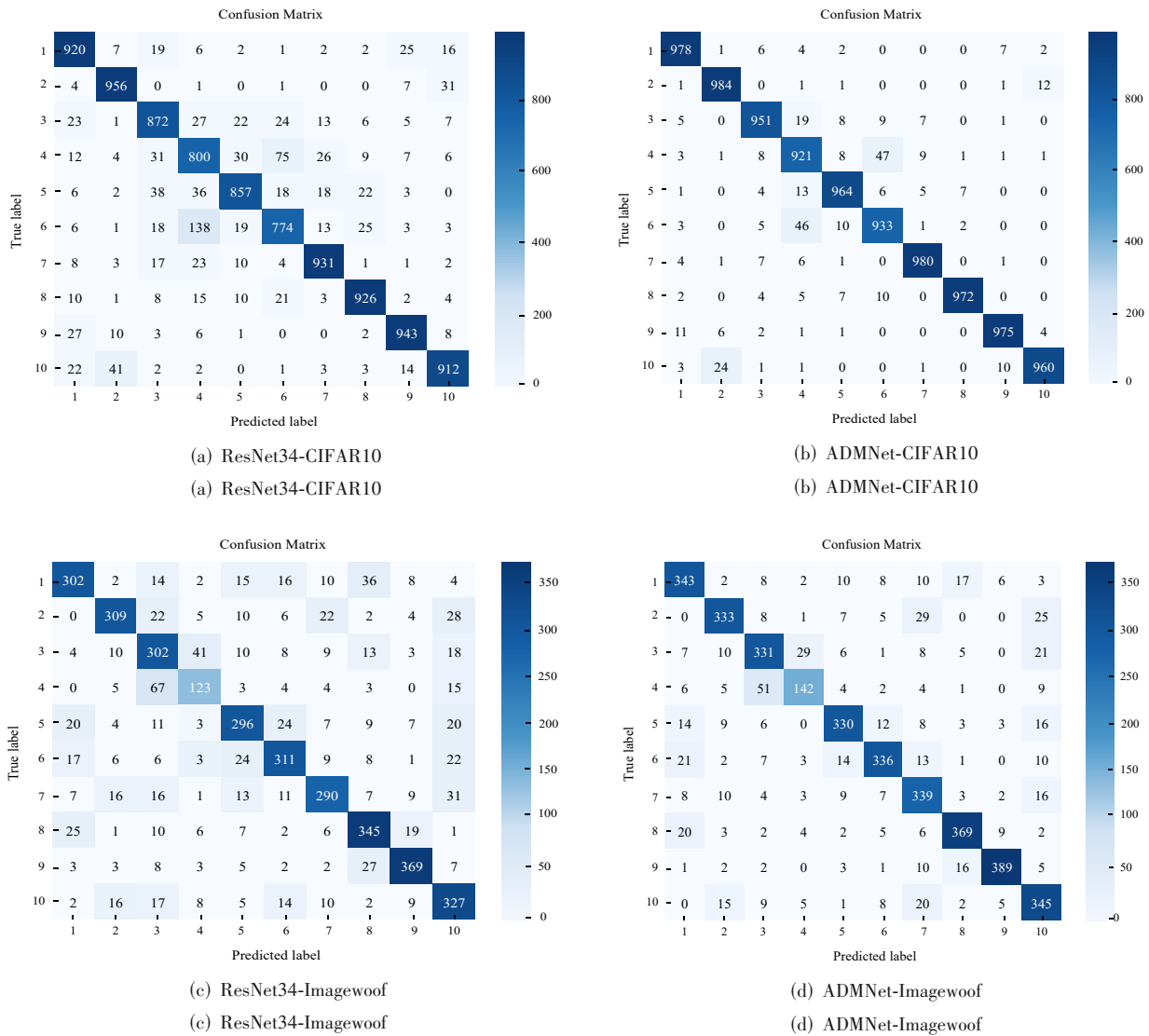


图 19 分类混淆矩阵

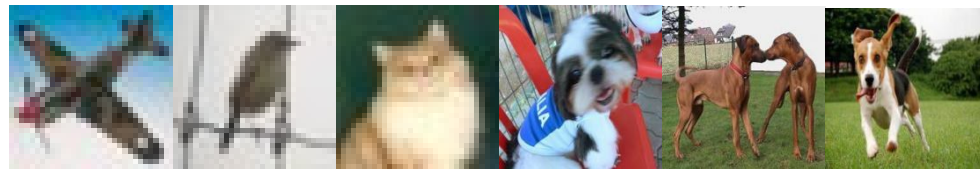
Figure 19 Classification confusion matrix

动策略^[31]对热力图进行定量评估。该策略的核心思想是对热力图中准确反映模型决策所依赖的关键区域逐步施加扰动(即逐步删除区域),并记录每次扰动后模型对原始真实类别的预测置信度,直至所有关键区域被遍历,再以累计被扰动的像素比例为横坐标、以模型对应的预测置信度为纵坐标,绘制置信度-扰动比例曲线,计算该曲线下的面积(Area Under the Curve, AUC)作为评价指标。AUC 越小,表明在扰动(删除)关键区域后该模型的预测置信度下降越快,反过来证明热力图所定位的区域对分类模型的决策至关重要,从而说明模型生成的热力图对关键特征区域的定位越准确。

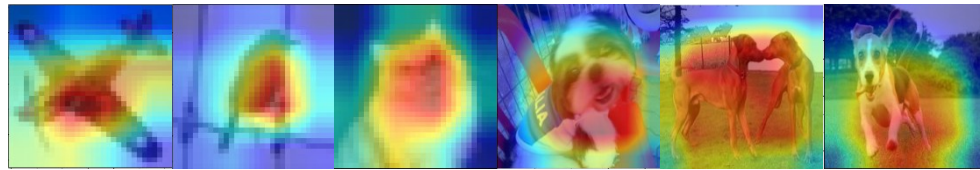
通过对本文网络与 ResNet-34、ShuffleNet、MobileNet_V2、VGGNet 和 GoogleNet 生成的图 20 中每一

列热力图进行 MoRF 扰动测试,得到 AUC 分数结果如表 4 所示。由表 4 数据可知,对 ADMNet 生成的热力图进行 MoRF 扰动测试后得到的 AUC 分数分别为 0.613 2、0.443 2、0.102 2、0.156 7、0.472 7 与 0.209 8,均低于其他五个网络,表明本文网络按照热力图指示的重要性顺序进行区域扰动时,模型预测置信度下降最迅速,印证其能够精准识别图像中影响分类决策的关键特征区域。

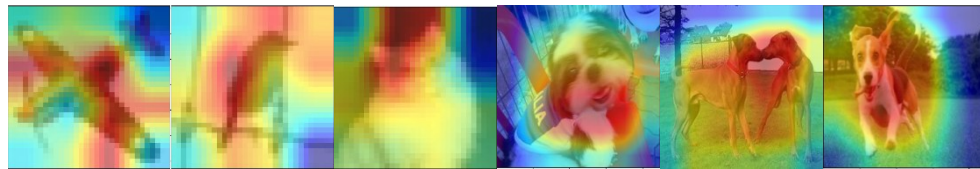
综上所述,本文所提出的特征衰减解耦模块能够有效降低图像关键特征与无关信息间的耦合度,增强关键特征的表达能力与判别性。此外,结合 FAP 模块的多尺度特征聚合功能,进一步保留更丰富的上下文信息,使热力图呈现出更符合生物认知的权重分布。相较于其他图像分类方法,本文所提出的 ADMNet 能够更精确地



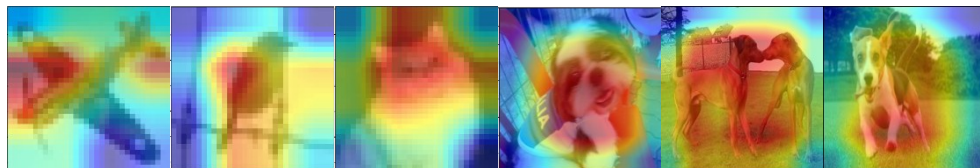
(a) 原始图像
(a) Original image



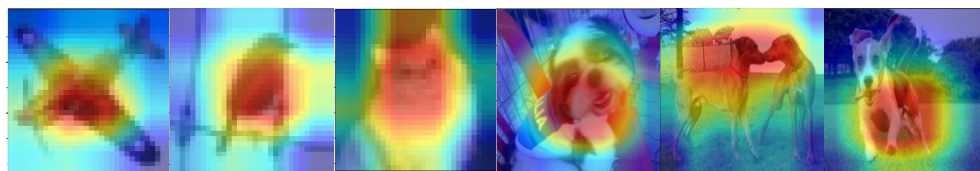
(b) ResNet-34
(b) ResNet-34



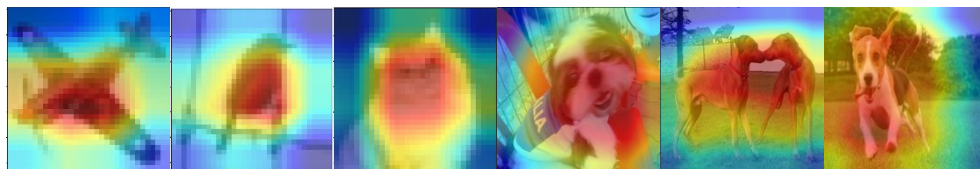
(c) ShuffleNet
(c) ShuffleNet



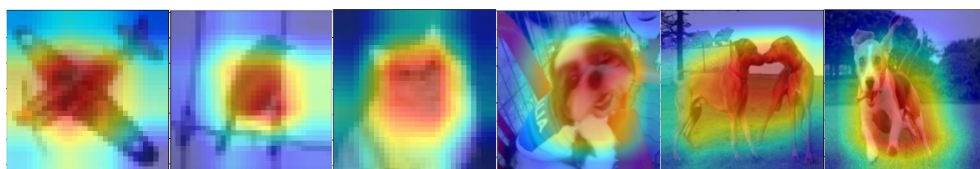
(d) MobileNet_V2
(d) MobileNet_V2



(e) VGGNet
(e) VGGNet



(f) GoogleNet
(f) GoogleNet



(g) ADMNet
(g) ADMNet

图 20 各网络生成的热力图

Figure 20 Heatmaps generated by each network

捕捉图像主体特征,从而提升网络分类精度。

表 4 各网络 MoRF 扰动测试下的 AUC 分数

Table 4 AUC scores under each network MoRF perturbation test

网络	第一列	第二列	第三列	第四列	第五列	第六列
ResNet-34	0.725 2	0.573 5	0.524 3	0.220 3	0.705 8	0.363 1
ShuffleNet	0.749 8	0.634 9	0.552 4	0.351 5	0.783 6	0.463 2
MobileNet_V2	0.890 4	0.658 9	0.617 9	0.295 7	0.763 0	0.479 9
VGGNet	0.740 5	0.550 6	0.338 7	0.220 4	0.649 8	0.382 2
GoogleNet	0.640 8	0.446 4	0.158 8	0.166 8	0.592 7	0.222 9
ADMNet	0.613 2	0.443 2	0.102 2	0.156 7	0.472 7	0.209 8

注:加粗字体表示各列最优结果。

4 结束语

针对现有图像分类网络提取到的特征耦合度高、缺乏判别性导致关键特征表达能力受限,进而影响分类精度的问题,本文提出衰减解耦机制的图像分类网络(ADMNet)。该网络的核心是特征衰减解耦(FAD)模块,该模块通过结合 SAD 与 CD 模块,分别从空间和通道维度对特征进行解耦。在空间维度,SAD 模块采用多阈值信号衰减函数,对图像特征的不同子空间进行差异化衰减变换,从而有效解耦各类特征表示。在通道维度,CD 模块利用多尺度一维卷积建模通道间的耦合关系,显著降低关键特征通道与无关通道的关联性与耦合度。上述模块通过双分支的联合操作,构成一个完整的衰减解耦机制,实现对图像特征的有效解耦,提高关键特征判别性,进而增强特征的非线性表达能力。此外,为减少网络传递过程中的信息损失,本文构建 FAP 模块,有效提升特征图的空间信息利用率。最后,通过调整残差块的内部结构,对特征表示进行优化,使 ADMNet 能够有效降低特征耦合度,增强关键特征的判别性,最终提高网络的图像分类能力。

本文验证了受生物神经元衰减机制启发的特征解耦方法在图像分类任务中的有效性与发展潜力。未来的研究将会聚焦于如何高效实现特征解耦,减少模型参数与计算复杂度,提高网络在其他任务领域的泛化性与迁移能力等问题,进一步提升网络的图像分类性能。

参考文献

- [1] Lecun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.
- [2] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6): 84-90.
- [3] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[PP/OL]. V6.arXiv (2015-04-10)[2025-12-20]. <https://arXiv.org/abs/1409.1556>.
- [4] Szegedy C, Liu Wei, Jia Yangqing, et al. Going deeper with convolutions[C]//Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2015: 7298594.
- [5] He Kaiming, Zhang Xiangyu, Ren Shaoqing, et al. Deep residual learning for image recognition[C]//Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2016: 770-778.
- [6] Khan S, Naseer M, Hayat M, et al. Transformers in vision: A survey[J]. ACM Computing Surveys, 2022, 54(10s): 1-41.
- [7] Touvron H, Cord M, Douze M, et al. Training data-efficient image transformers & distillation through attention[C]//Proceedings of the 38th International Conference on Machine Learning. PMLR, 2021: 10347-10357.
- [8] Liu Ze, Lin Yutong, Cao Yue, et al. Swin transformer: Hierarchical vision transformer using shifted windows[C]//Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE, 2022: 9992-10002.
- [9] Xu Wanlu, Liu Hong, Shi Wei, et al. Adversarial feature disentanglement for long-term person re-identification[C]//Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence. California: IJCAI.org, 2021: 1201-1207.
- [10] Liu Dongnan, Zhang Chaoyi, Song Yang, et al. Decompose to adapt: Cross-domain object detection via feature disentanglement[J]. IEEE Transactions on Multimedia, 2023, 25: 1333-1344.
- [11] 唐珺琨, 张辉, 张邹铨, 等. 基于任务关联特征解耦网络的无监督领域自适应图像分类[J]. 计算机科学, 2023, 50(增刊2): 293-300.
Tang Junkun, Zhang Hui, Zhang Zouquan, et al. Image classification for unsupervised domain adaptation based on task relevant feature separation network[J]. Computer Science, 2023, 50(S2): 293-300. (in Chinese)
- [12] Yuan Jianjun, Wu Fujun, Zhao Luoming, et al. MLFD: Multi-level feature disentanglement robust network with a novel maximum-minimum algorithm for medical image recognition[J]. Expert Systems with Applications, 2025, 294: 128815.
- [13] Jiang Wentao, Yuan Heng, Liu Wanjun. Neuron signal attenuation activation mechanism for deep learning[J]. Patterns, 2025, 6(1): 101117.
- [14] Chen L C, Papandreou G, Schroff F, et al. Rethinking atrous convolution for semantic image segmentation[PP/OL]. V3. arXiv (2017-12-05)[2025-12-20]. <https://arXiv.org/abs/1706.05587>.
- [15] Tan Mingxing, Le Q V. EfficientNet: Rethinking model scaling for convolutional neural networks[C]//Proceedings of the 36th International Conference on Machine Learning. Long Beach: PMLR, 2019: 6105-6114.
- [16] Han Kai, Wang Yunhe, Tian Qi, et al. GhostNet: More features from cheap operations[C]//Proceedings of the

- 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2020: 1577-1586.
- [17] Huang Gao, Liu Zhuang, Van Der Maaten L, et al. Densely connected convolutional networks[C]//Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2017: 2261-2269.
- [18] Zhou Chenlin, Zhang Han, Zhou Zhaokun, et al. QKFormer: Hierarchical spiking transformer using Q-K attention[C]//Proceedings of the 38th International Conference on Neural Information Processing Systems. New York: Curran Associates Inc., 2024: 13074-13098.
- [19] Zagoruyko S, Komodakis N. Wide residual networks[PP/OL]. V4. arXiv (2017-06-14)[2025-12-20]. <https://arXiv.org/abs/1605.07146>.
- [20] Lan Hai, Wang Xihao, Shen Hao, et al. Couplformer: Rethinking vision transformer with coupling attention[C]//Proceedings of the 2023 IEEE/CVF Winter Conference on Applications of Computer Vision. Piscataway: IEEE, 2023: 6464-6473.
- [21] Hasanpour S H, Rouhani M, Fayyaz M, et al. Towards principled design of deep convolutional networks: Introducing SimpNet[PP/OL]. V1. arXiv (2018-02-17)[2025-12-20]. <https://arXiv.org/abs/1802.06205>.
- [22] Qiu Xuerui, Zhu Ruijie, Yuhong Chou, et al. Gated attention coding for training high-performance and efficient spiking neural networks[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2024, 38(1): 601-610.
- [23] Konstantinidis D, Papastratis I, Dimitropoulos K, et al. Multi-manifold attention for vision transformers[J]. IEEE Access, 2023, 11: 123433-123444.
- [24] Shin H, Choi D W. Teacher as a lenient expert: Teacher-agnostic data-free knowledge distillation[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2024, 38(13): 14991-14999.
- [25] Ma Chenxiang, Wu Jibin, Si Chenyang, et al. Scaling supervised local learning with augmented auxiliary networks[C]//Proceedings of the Twelfth International Conference on Learning Representations. OpenReview.net, 2024. <https://arxiv.org/pdf/2402.17318>
- [26] Choromanski K M, Likhoshesterov V, Dohan D, et al. Rethinking attention with performers[C]//Proceedings of the 9th International Conference on Learning Representations. OpenReview.net, 2021.
- [27] Wu Xidong, Gao Shangqian, Zhang Zeyu, et al. Auto-train-once: Controller network guided automatic network pruning from scratch[C]//Proceedings of the 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2024: 16163-16173.
- [28] Yuan Heng, Liu Jie, Jiang Wentao, et al. Double-pooling residual classification network based on feature reordering attention mechanism[J]. Journal of Image and Graphics, 2025, 30(1): 110-129.
- [29] Zhang Xiangyu, Zhou Xinyu, Lin Mengxiao, et al. ShuffleNet: An extremely efficient convolutional neural network for mobile devices[C]//Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 6848-6856.
- [30] Sandler M, Howard A, Zhu Menglong, et al. MobileNetV2: Inverted residuals and linear bottlenecks[C]//Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 4510-4520.
- [31] Samek W, Binder A, Montavon G, et al. Evaluating the visualization of what a deep neural network has learned[J]. IEEE Transactions on Neural Networks and Learning Systems, 2017, 28(11): 2660-2673.

作者简介



袁 姮 女,1988年2月出生于湖北省黄冈市。现为辽宁工程技术大学软件学院副教授。主要研究方向为图像与视觉信息计算、模式识别与人工智能。

E-mail: Intuyuanheng@163.com



杨继真 女,2002年7月出生于河南省南阳市。现为辽宁工程技术大学软件学院硕士研究生。主要研究方向为图像与视觉信息计算、模式识别与人工智能。

E-mail: 2444603170@qq.com



张晟翀 男,1973年4月出生于安徽省合肥市。高级工程师。主要研究方向为数字信号处理、模式识别与人工智能。

E-mail: zsc417@126.com