

# 融合渐进训练策略的 logo 图像分类<sup>①</sup>



麻宇轩, 齐亚莉

(北京印刷学院 信息工程学院, 北京 102600)

通信作者: 麻宇轩, E-mail: 18801252926@163.com

**摘要:** 经济全球化赋予了 logo 巨大的商业价值, 随着计算机视觉领域的发展, 为 logo 分类与识别提供了更广阔的应用领域. 本文针对 logo 图像的分类识别, 为了提高模型对 logo 图像分类的能力, 基于 logo 图像整体特征不显著且数量众多的特点, 提出了用细粒度图像分类的方法渐进式多粒度拼图训练 (progressive multi-granularity training of jigsaw patches, PMG-Net) 对 logo 图像数据集进行分类. 通过拼图生成器生成包含不同粒度信息的输入图像, 再引入渐进式多粒度训练模块融合不同粒度的特征, 融合后的特征更注重图像之间的细微差别, 使 logo 图像分类的效果有显著提高. 在提取输入图像特征时采用 LeakyReLU (leaky rectified linear unit) 激活函数保留图像中的负值特征信息, 并引入通道注意力机制, 调整特征通道的权重, 增强特征信息指导能力以改进模型分类效果. 实验结果表明, 本文在 logo 图像数据集上的分类精确率优于传统的分类方法. 本文通过融合多粒度特征的渐进训练策略以及随机拼图生成器的方法实现了对 logo 图像的高效分类, 为解决 logo 图像分类中存在的问题提供了一种新的思路.

**关键词:** logo 图像; 细粒度图像分类; 渐进训练策略; LeakyReLU; 注意力机制

引用格式: 麻宇轩, 齐亚莉. 融合渐进训练策略的 logo 图像分类. 计算机系统应用, 2023, 32(6): 130–139. <http://www.c-s-a.org.cn/1003-3254/9069.html>

## Logo Image Classification Incorporating Progressive Training Strategies

MA Yu-Xuan, QI Ya-Li

(School of Information Engineering, Beijing Institute of Graphic Communication, Beijing 102600, China)

**Abstract:** Economic globalization has given logo a huge commercial value, and the development of the computer vision provides a broader application field for logo classification and recognition. This study considers that the overall features of logo images are not significant, and the number of images is large, and then it proposes progressive multi-granularity training of jigsaw patches (PMG-Net), a method of fine-grained image classification, to classify the logo image dataset, so as to improve the ability of the model to classify logo images. The input images containing different granularity information are generated by the jigsaw patch generator, and then the progressive multi-granularity training module is introduced to fuse the features of different granularities. The fused features pay more attention to the subtle differences between images so that the effect of logo image classification is significantly improved. The leaky rectified linear unit (LeakyReLU) activation function is used to retain the negative feature information in the image when the input image features are extracted, and the channel attention mechanism is introduced to adjust the weights of the feature channels, so as to enhance the feature information guiding ability and improve the classification effect of the model. The experimental results show that the classification accuracy of this study on the logo image dataset is better than that of traditional classification methods. This study achieves efficient classification of logo images by incorporating a progressive training strategy with multi-granularity features and a random jigsaw patch generator, which provides a new idea to solve the

<sup>①</sup> 基金项目: 北京印刷学院科研创新团队项目 (20190122019)

收稿时间: 2022-09-09; 修改时间: 2022-10-10; 采用时间: 2022-11-29; csa 在线出版时间: 2023-04-25

CNKI 网络首发时间: 2023-04-26

existing problems in logo image classification.

**Key words:** logo image; fine-grained image classification; progressive training strategy; leaky rectified linear unit (LeakyReLU); attention mechanism

## 1 引言

Logo 是企业、组织或者是机构的一个可视化标志,它由字形、文本信息和图像组合而成,可以代表一个组织或是机构背后的品牌的理念与独特性。Logo 的分类与识别在广泛的应用中非常重要。最近很多研究主要针对在自然场景下商标的相关任务,尤其在产品与品牌识别、版权保护、品牌溯源追踪、商品图像搜索、商业广告分析与智能交通等领域都有着巨大的应用前景。在品牌识别、溯源追踪与版权保护领域,构建完善的电子商品目录,帮助企业快速检测出相关品牌,完成相应品牌的商品个性化介绍与推荐和盗版侵权的检测与追究。在商业广告领域,基于 logo 的品牌搜索,品牌的广告投放以及社交媒体上的受欢迎程度和发展趋势,都让 logo 具有巨大的商业价值。对公文标志进行识别分类,可以提高文件的处理效率。在智能交通领域,路标 logo 检测技术的研究是智能交通系统的重要组成部分,现代交通汽车数量骤增、安全事故频发,智能交通系统的出现十分必要,因此 logo 的分类研究具有重大的意义。

传统分类方法上,logo 识别是通过基于关键点的检测器和描述符来解决的。Revaud 等人认为解决 logo 识别这一难题主要在于是 logo 与其他随机图像中常见的局部图案非常相似,针对这一难题提出通过学习一个统计模型来理解图像匹配算法所输出的不正确检测的分布,从而改善匹配标准<sup>[1]</sup>。Boia 等人提出了一种解决自然图像中的品牌 logo 定位和分类问题的新方法<sup>[2]</sup>。通过将各个训练图像映射到中心图像上构建了一个类模型,利用这个模型进行标志的定位和识别,再利用反向二级模型来处理反向颜色实例。Kumar 等人提出了一种基于 logo 外观的 logo 分类系统图像<sup>[3]</sup>,利用 logo 图像的全局特征,如颜色、纹理、形状等进行分类。然而,现有的大部分分类工作的成功,包括采用传统模式识别算法的 logo 分类,主要取决于所选择的特征类别。近期的工作中,深度学习技术由于能够提取抽象层次更高的特征而得到快速发展。特别是,使用卷积神经网络模型的方法是计算机视觉任务中取得了显著提高的

结果。Bianco 等人提出了一种利用深度学习专门针对 logo 识别分类的管道<sup>[4]</sup>,该管道由 logo 区域建议和一个专门用于分类的 CNN 组成。Hou 等人展开了更加深入的研究<sup>[5]</sup>,提出了一种算法结合卷积神经网络和传统模式识别算法实现数据的深度特征表示。

与一般的图像分类任务相比,logo 图像分类任务中 logo 图像的整体特征并不显著且品牌数量众多,导致类间差异过小以及无法准确捕捉到图像显著特征,分类模型需要可以达到准确区分目标图像关键部位特征的能力。所以本文使用细粒度图像分类问题来解决这一难题。

针对上述问题,本文提出了一种改进的 PMG-Net 的 logo 图像分类方法,本文模型利用 PMG-Net (progressive multi-granularity training of jigsaw patches)<sup>[6]</sup>作为基本框架,在提取图像特征时引入 LeakyReLU (leaky rectified linear unit) 激活函数保留图像中的负值信息,防止训练时的神经元失效,并引入通道注意力机制,提高模型的特征提取能力,最终实现对 logo 图像的分类。本文在 Medical 数据集和 Transportation 数据集两大数据集上实验,通过大量实验证明了方法的有效性。

## 2 相关技术

PMG-Net 网络分类模型是 Du 等人<sup>[6]</sup>在 2020 年提出的一种细粒度图像分类的方法。网络模型分成两部分,分别是拼图生成器和渐进训练阶段。拼图生成器 (jigsaw puzzle generator) 将原图均匀的划分并随机打乱,生成新的拼图图像。渐进训练阶段是建立多粒度渐进训练框架用于学习图像不同粒度之间的互补信息。PMG-Net 是将上述两个模型加以结合,形成一个渐进式训练网络模型,可以得到不同粒度级别的特征,使分类网络能够关注局部区域的特征,达到提高分类精度的目的,有效地对细粒度图像进行分类。

### 2.1 网络模型概述

本文提出的改进的 PMG-Net 网络模型主要分为 3 个步骤。首先采用拼图生成器,将完整图像在不同阶段分割成不同尺度的拼图块后进行打乱生成新的

拼图图像;随后采用 ResNet 网络作为特征提取骨干网络,并针对  $ReLU$  激活函数在输入为负的区段梯度恒为 0 的问题将  $ReLU$  激活函数改进为  $LeakyReLU$  激活函数.同时引入 SE 通道注意力模块,帮助网络模型可以有选择性地加强对分类有用的特征且抑制对分类无用的特征;最后,通过渐进式训练从低阶段开

始训练,逐步增加新的训练阶段,由于不同阶段的输入图像的粒度不同,这样可以保证在每个阶段都能专注于学习自己对应阶段粒度的特征.在较低阶段,感受野和表示能力有限,网络学习局部区域的辨别性信息;在较高阶段,更大的感受野和更强的表达能力使网络能够学习全局结构.本文网络模型如图 1 所示.

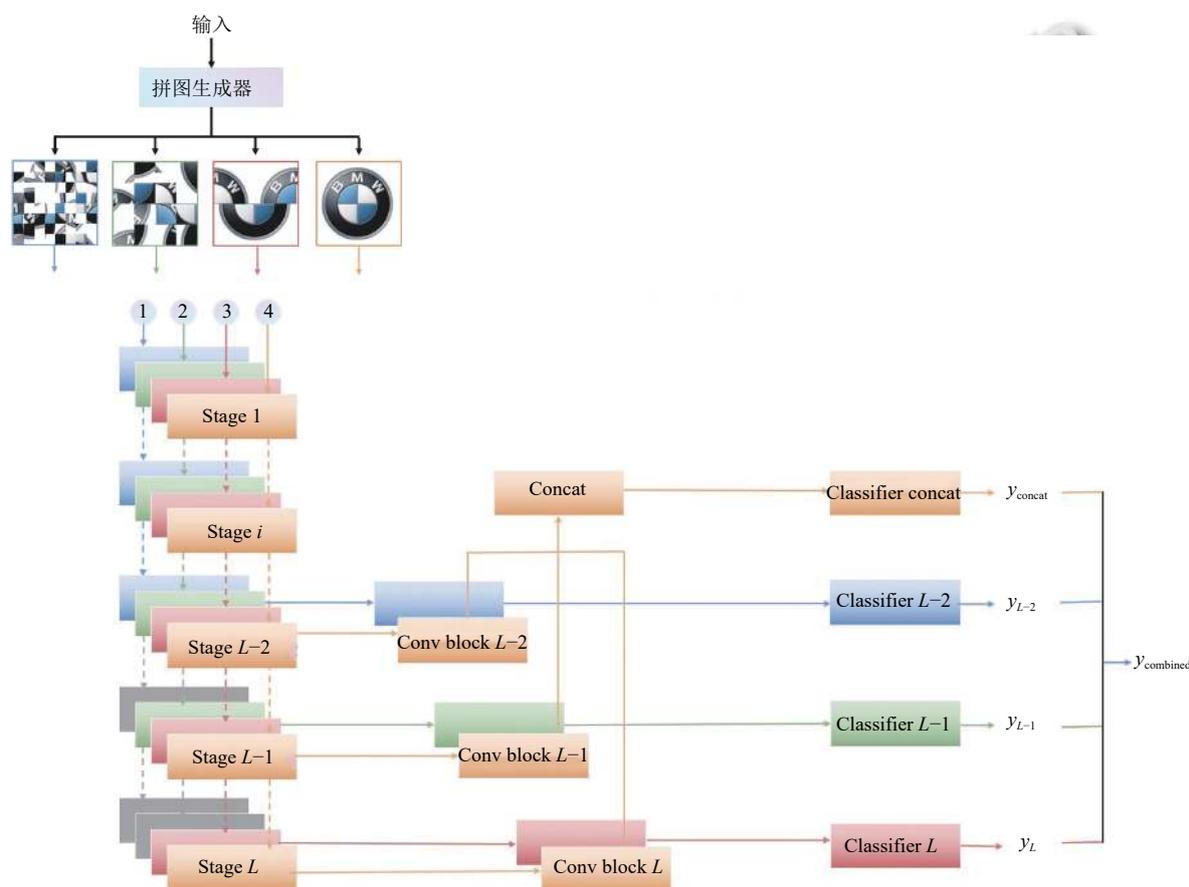


图 1 网络结构图

## 2.2 拼图生成器

Wei 等人提出了一种适用于任意网格尺寸与维度的“拼图”问题的新方法<sup>[7]</sup>.他们将无标签的图像分割成图像块并打乱图像块的顺序,以迭代的方式逐步调整图像块的顺序直到收敛,用于特征学习并验证了其有效性.本文引入了拼图的概念,提出拼图生成器.拼图生成器模块借助拼图游戏的思想,将完整图像在不同阶段分割成不同尺度的拼图块后进行打乱生成新的拼图图像,为渐进式训练阶段生成输入图像.

给定输入图像的大小为  $P \in R^{W \times H}$ , 将它分成  $n \times n$

个大小相等的拼图块,拼图块的大小为  $W/n, H/n$ , 这里必须满足可以同时整除  $W$  和  $H$ .然后将拼图块随机打乱顺序后生成一个新的图像  $P' \in R^{W \times H}$ .在训练过程中,拼图生成器将完整的输入图像分割成拼图块,随着步骤的进行,拼图块的大小随之增加,直到最后一步输入原始图像.拼图块大小的不同反映了粒度的等级,设计不同粒度的图像,迫使网络模型在不同的训练阶段只能学习当前等级的粒度信息.

在选择超参数  $n$  的值时满足下面几个条件.拼图块的大小要小于对应阶段的感受野,否则会降低拼图生

成器的性能;二是随着各阶段感受野的增加,拼图块的大小应成比例增加,每个阶段通常为上一个阶段的两倍。

在将输入的完整图像裁剪成拼图块时,并不能保证所有小于拼图块的部分是完整的,有时可能会被拆分,例如 logo 图像中的字母显示不完整,但是这并不会影响模型的性能。本文的网络结构提出的数据增强策略是随机裁剪策略,所以即便是同一张图像,每次训练的角度也都不相同,这样也就保证了在特定粒度等级上能获取更多的有区别的细节。

### 2.3 渐进训练网络

由于 logo 图像的种类多种多样,随着时间的变化,部分企业或品牌 logo 图像会出现更新换代的情况。为了解决这种较大的类内变化,提出使用这种较浅层次中学习稳定的细粒度信息,并随着训练的进行逐渐将注意力转移到较深的层中学习大粒度级别的抽象信息的方法,也就是渐进式训练方法。

渐进训练被广泛应用于生成任务,最初是 Karras 等人提出的一种生成高分辨率图像的方法 PGGAN (progressive growing of GANs)<sup>[8]</sup> 中提到的。后来的工作中,PCRN (progressive cascading residual network)<sup>[9]</sup>, ProSR (progressive super-resolution networks)<sup>[10]</sup>, StyleGAN (style-based generative adversarial networks)<sup>[11]</sup>, SinGAN (single image generative adversarial networks)<sup>[12]</sup> 都使用了渐进式训练生成图像。它们从低分辨率的图像开始训练生成器和判别器,向网络中添加新的层来逐步提高分辨率。PMG-Net<sup>[6]</sup> 提出了一种用于细粒度图像分类的渐进式训练方法,该方法通过前一阶段的训练指导下一个阶段的训练,逐渐将重点从局部特征转移到全局特征。本文使用的渐进式训练方法在 PMG-Net 的基础上引入了通道注意力模块,更加充分地利用了模型的特征提取能力,从而使模型的分类性能得到显著提高。

渐进式训练策略的目标是提取不同粒度的辨别信息 (discriminative information)。“渐进式”指的是先训练低级阶段,然后逐步增加新的训练阶段。模型的特征逐渐被送到更高阶段,网络对于区别性信息的定位也是从局部到逐渐放大到全局结构,并不是同时学习所有粒度的信息。本文用  $F$  表示主干特征提取器,它有  $S$  个阶段,抽取  $l$  个阶段输出的特征图  $S_l$  用于渐进训练。同时引入了卷积块  $B_l^{\text{conv}}$ ,将  $S_l$  传入预设好的卷积模块并将其简化为向量表示  $V_l = B_l^{\text{conv}}(S_l)$ 。为了预测概率分布,

定义分类模块为  $B_l^{\text{class}}$ ,将不同的向量传入不同的分类模块,得到不同粒度的预测概率。将所有的特征向量  $V_l$  合并输入到分类模块中,得到融合所有粒度信息的概率分布  $y_l = B_l^{\text{class}}(V_l)$ 。如图 1 步骤一步一步说明了训练过程。

渐进式训练可以在不同的中间阶段施加分类损失。对于每个阶段的输出和连接特征的输出的训练,模块采用交叉熵损失  $L_{\text{CE}}$  来计算真实标签和预测概率分布之间的距离。可通过以下方式计算:

$$L_{\text{CE}}(y^l, y) = - \sum_{i=1}^m y_i^l \times \log(y_i^l) \quad (1)$$

其中,  $m$  代表的是类别的数量,  $y_i^l$  代表的是输入  $X$  在  $l$  阶段属于第  $i$  个类别的概率。

为了提高网络模型的性能,将最后多个阶段的输出串联为:

$$V_{\text{concat}} = \text{concat}[V_{L-S+1}, \dots, V_{L-1}, V_L] \quad (2)$$

其中,  $S$  表示最后阶段的数量。接下来是一个分类模型,  $y_{\text{concat}} = H_{\text{concat}}^{\text{class}}(V_{\text{concat}})$ 。然后,我们使用交叉熵损失函数优化:

$$L_{\text{CE}}(y_{\text{concat}}, y) = - \sum_{i=1}^m y_i^i \times \log(y_{\text{concat}}^i) \quad (3)$$

在测试阶段,如果只利用  $y_{\text{concat}}$  作为预测,那么就是只输入完整的图像进入训练模型,模块的最终结果可以表示为:

$$P_{\text{concat}} = \arg \max(y_{\text{concat}}) \quad (4)$$

但是各个阶段的预测是互补的,将不同粒度信息得到输出组合到一起,可以得到更好的性能和结果,得到最终的结果可以表示为:

$$P_{\text{mix}} = \arg \max \left( \sum_{l=L-S+1}^L y_l + y_{\text{concat}} \right) \quad (5)$$

### 2.4 激活函数改进

传统的 ResNet 网络一般使用 ReLU 激活函数,用以提高网络的非线性表达能力和训练速度。ReLU 激活函数的公式为:

$$\text{ReLU}(x) = \begin{cases} x, & x > 0 \\ 0, & x \leq 0 \end{cases} \quad (6)$$

ReLU 函数曲线如图 2 所示。ReLU 激活函数在输入为负的区域梯度恒为 0,梯度得不到更新,神经

元不会被任何数据激活,这种现象称为神经元的“坏死”,导致网络在训练的过程中无法提取特征.为了解决此问题,本文网络模型在选用 ResNet 网络作为特征提取网络时使用 LeakyReLU 激活函数替换原网络中的 ReLU 激活函数,LeakyReLU 激活函数的公式如式(7)所示,其中 $\alpha$ 表示偏移量.

$$LeakyReLU(x) = \begin{cases} x, & x > 0 \\ \alpha x, & x \leq 0 \end{cases} \quad (7)$$

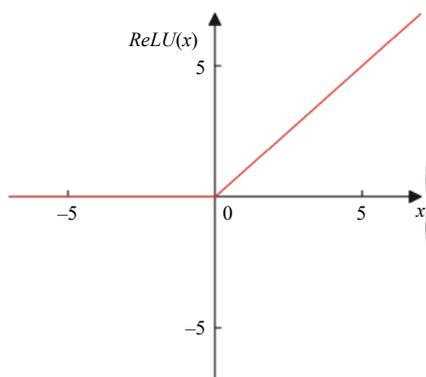


图2 ReLU 函数曲线

LeakyReLU 函数图像如图3所示.与 ReLU 函数相比,LeakyReLU 函数保留了 ReLU 函数正区间的特点,又在负区间设置了较小的斜率,保证了在输入值为负时的输出不为零,从而解决了 ReLU 函数在输入为

负区间神经元不学习的问题.改进后的网络保留了有用的负值特征,保证了梯度的更新,增强了网络模型的表达能力,提高了分类的准确率.

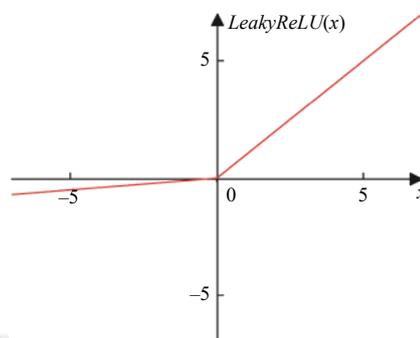


图3 LeakyReLU 函数曲线

### 2.5 SE-Net 通道注意力机制

SENet (squeeze-and-excitation networks) 采用一种“特征重标定”的方法,其核心在于 SE 模块.具体来说,就是通过学习的方式来自动获取到每个特征通道的重要程度,然后依照这个重要程度去提升有用的特征,抑制对当前任务用处不大的特征<sup>[13]</sup>,这样的特点可以帮助网络模型更好地提取对于图像分类任务的有效特征.为了使模型有选择性地加强对分类有用的特征且抑制对分类无用的特征,本文在 ResNet 网络的基础上引入了 SE 通道注意力模块,模块的结构如图4所示.

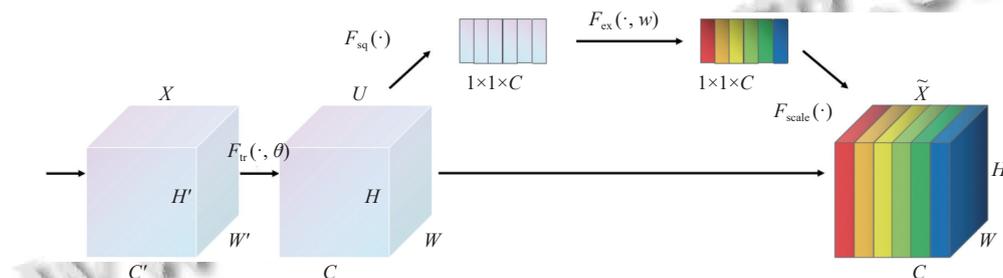


图4 SENet 通道注意力模块图

SE 模块主要由 3 部分组成,分别是压缩 (squeeze) 操作,激励 (excitation) 操作和重标定 (reweight) 操作.首先是 squeeze 操作对输入大小为 $H \times W \times C$ 的特征图采用全局平均池化进行特征压缩,得到 $1 \times 1 \times C$ 的特征向量. Squeeze 操作的表达式如式(8)所示,其中, $F_{sq}(\cdot)$ 为 squeeze 操作,为经过特征压缩后得到的特征向量.

$$Z_c = F_{sq}(U_c) = \frac{1}{W \times H} \sum_{i=1}^W \sum_{j=1}^H U_c(i, j) \quad (8)$$

然后 excitation 操作通过两个全连接层学习通道间的相关性并得到特征通道的权重,大小为 $1 \times 1 \times C$ .表达式如式(9)所示.其中, $F_{ex}(\cdot)$ 为 excitation 操作, $S$ 为得到的特征通道的权重, $\sigma$ 为 Sigmoid 激活函数,

$\delta$ 为 ReLU 激活函数,  $W_1$ 和 $W_2$ 分别表示两个全连接层的权重.

$$S = F_{\text{ex}}(Z, W) = \sigma(W_2\delta(W_1Z_c)) \quad (9)$$

这一系列操作后可以获得各个通道的重要程度, 然后按照重要程度的顺序促进有用的特征并抑制无用的特征.

本文选择在 ResNet 网络中嵌入 SENet 模块, 具体的网络结构如图 5 所示. 基于通道注意力机制的方法, 改进后的网络模型比仅采用一个全连接层次的 ResNet 网络更好地关注有效的分类信息, 更好地拟合通道之间的复杂性, 在计算复杂度没有显著增加的情况下, 提高分类的准确率, 获得更好的性能.

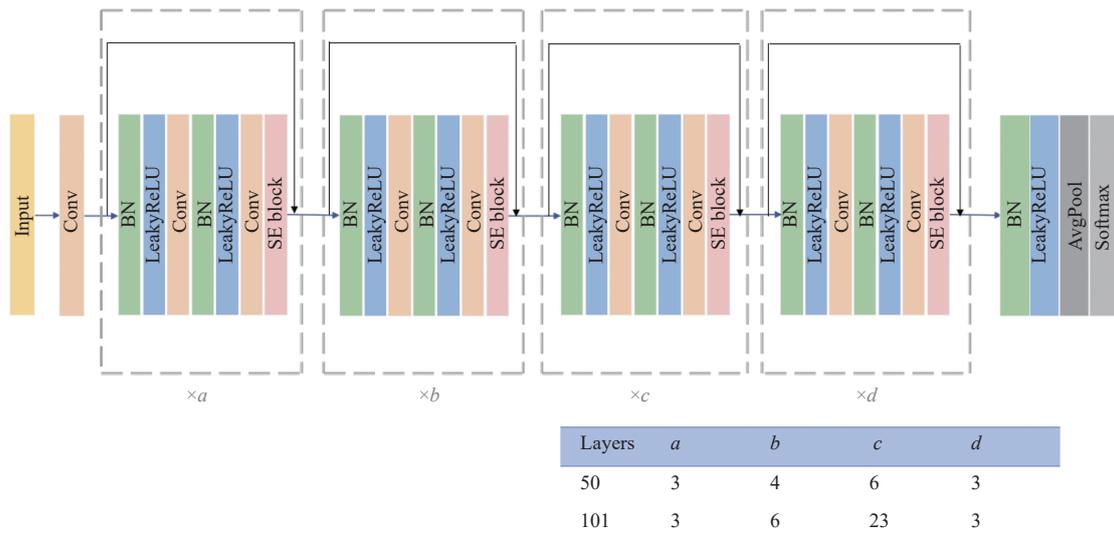


图 5 嵌入 SENet 模块的 ResNet 体系结构

### 3 实验结果与分析

#### 3.1 数据集

本文建立了两个新的 logo 数据集. 由于 COVID-19 的大流行引起的全球关注给医疗保健行业带来了巨大挑战, 为了给医疗行业提供更便利的帮助, 本文建立了包含医疗品牌 logo 的 Medical 数据集. 同样为了智能交通的完善, 建立了包含各种交通相关品牌的 Transportation 数据集.

本文使用的自构建的数据集分别是 Medical 数据集和 Transportation 数据集对分类网络进行评测, 以验证本文提出的方法在数据集上的有效性.

##### (1) Medical 数据集

由 2 006 张图像组成, 包含 28 个 logo 种类. 其中训练集 1 374 张图片, 测试集 632 张图片. 图 6 报告展示了 Medical 数据集中每个类的 logo 示例图像.

##### (2) Transportation 数据集

由 6 894 张图像组成, 显示 94 个 logo 种类. 其中训练集 4 888 张照片, 测试集 2 006 张照片. 图 7 展示了 Transportation 数据集中部分类别的 logo 示例图像,

图 8 选取了数据集中部分类别, 展示了同一种类 logo 的不同图像, 每一行展示一种类别.

#### 3.2 实验装置

本文实验是基于 PyTorch 框架上实现的. 实验使用 ResNet-101 网络为特征提取器, 将输入的图像大小调整到  $550 \times 550$ , 随机裁剪为  $448 \times 448$ , 进行随机水平反转. 网络模型使用动量梯度下降优化和批标准化作为正则化, 总阶段设置  $L=5, S=3$ . 模型的初始学习率设置为 0.002, 且在训练过程中引入余弦退火 (cosine annealing) 的学习率下降方式. 预训练的卷积层的学习率保持为新添加层的 1/10. 对于上述所有模型, 我们对它们进行了 100 个 epoch 的训练, 批大小为 16, 使用的权重衰减为 0.0005, 动量为 0.9.

#### 3.3 实验验证与分析

为了验证网络模型 PMG-Net 对 logo 图像分类的有效性, 在 Medical 数据集和 Transportation 数据集上进行了实验. Medical 数据集的准确率达到 72.47%, Transportation 数据集的准确率达到 77.71%. Medical 数据集的准确率图和损失率图如图 9 所示, Transpor-

tation 数据集如图 10 所示。

为了验证本文的网络模型对 logo 图像分类的有效性, 将 Transportation 数据集在 AlexNet、VGG-16、ResNet-50、PMG-Net 和本文网络模型上进行分类, 记录其结果作为对比数据, 实验结果如表 1 所示。本文的网络模型在分类上取得的效果最好, 能够获得 77.71% 的准确率, 相较于 AlexNet、VGG-16、ResNet-50、

PMG-Net 等网络模型, 准确率分别提高了 28.91%、14.88%、11.37%、1.37%。图 11 展示了各结果的对比, 根据实验结果对比可以看出, 本文的网络模型在 Transportation 数据集中, 比以往的方法有着更好的表现, 这说明通过采用 LeakyReLU 激活函数代替 ReLU 激活函数, 以及引入 SENet 通道注意模块, 可以很大程度上的提高模型的分类精度。



图 6 Medical 数据集内部分品牌 logo 图片



图 7 Transportation 数据集部分 logo 品牌

### 3.4 消融实验

(1) 为了验证在 ResNet 网络层数对于特征提取效果的影响, 本文设计了分别使用 ResNet-50 与 ResNet-101 网络作为特征提取网络的 PMG-Net 对两个数据集分

类的对比实验, 实验结果如表 2 所示。本文网络模型以 ResNet-50 和 ResNet-101 为基础模型都获得了非常先进的性能。实验结果表明, 在 Medical 数据集上, ResNet-50 网络的准确率达到 70.62%, ResNet-101 网络的准确率

为 72.47%。ResNet-101 比 ResNet-50 的精度提高了接近 2%。在 Transportation 数据集上, ResNet-50 网络的准确率达到 76.92%, 而 ResNet-101 网络达到 77.71%, 提高了 0.79% 的精度。综合两个数据集都可以看出以 ResNet-101 为基础模型作为本次实验的特征提取网络表现更好, 模型性能得到了明显的提升。各网络的正确

率对比如图 12 所示。

(2) 为了验证拼图生成器的有效性, 在推理阶段不再使用拼图生成器仅输入原始图像到训练模型中, 也就是在各个阶段都使用原始图像进行训练, 这里得到的正确率用 Accuracy\_a 表示。与正常参与本文模型训练的正确率 Accuracy 的比较如表 3 所示。

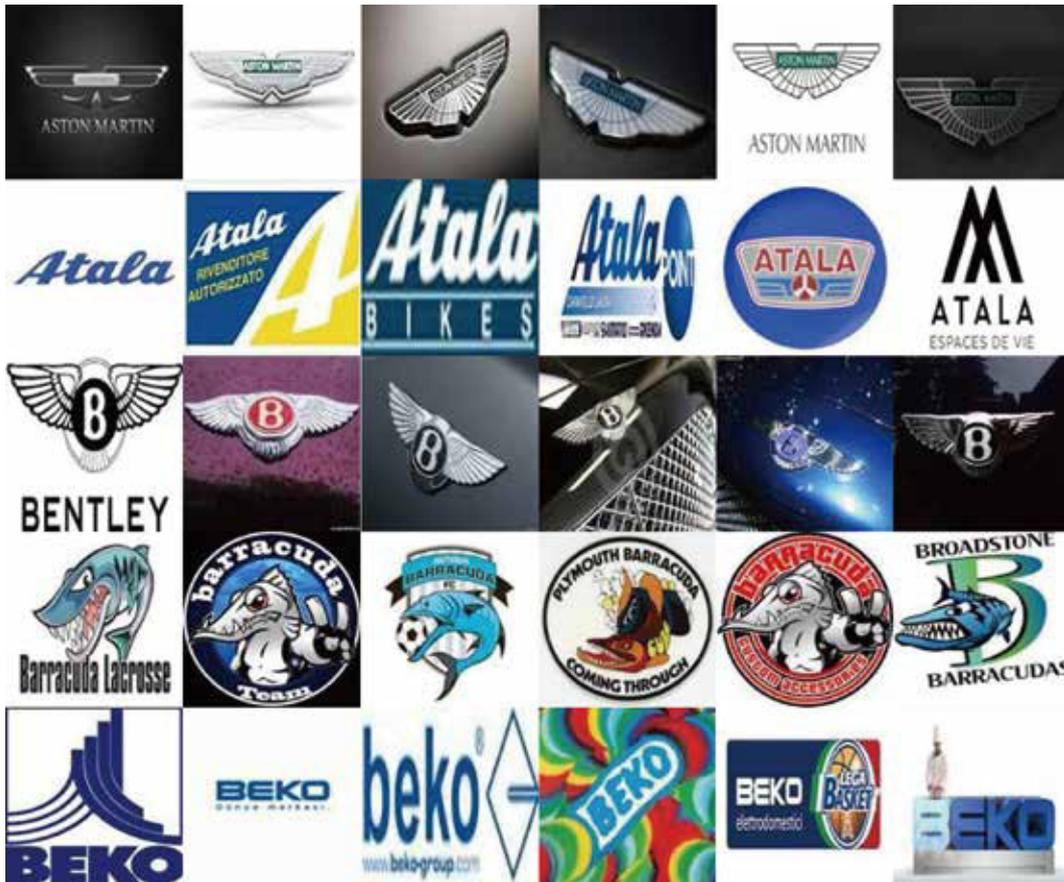


图 8 Transportation 数据集中同一品牌 logo 图片的不同照片

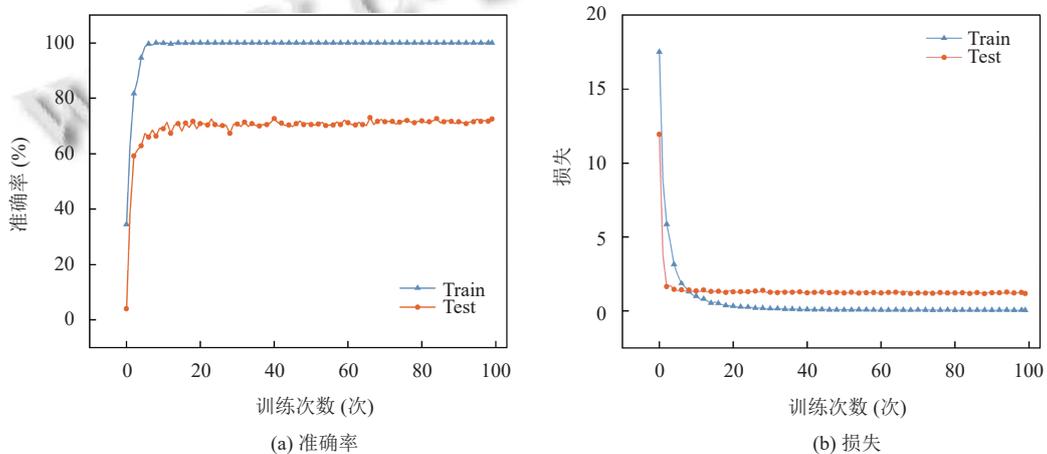


图 9 Medical 数据集的准确率和损失

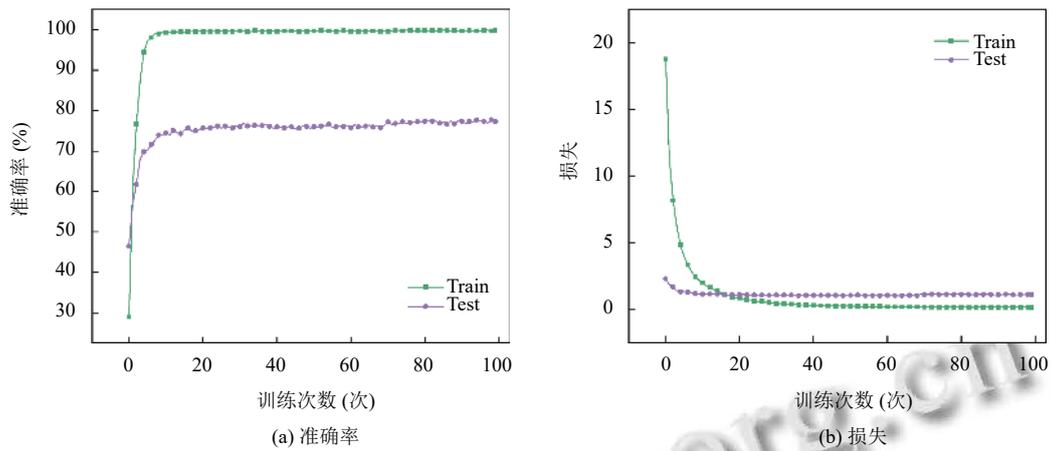


图 10 Transportation 数据集的准确率和损失

表 1 Transportation 数据集分类结果 (%)

方法	准确率
AlexNet	48.80
VGGNet-16	62.83
ResNet-50	66.34
PMG-Net	76.34
本文	77.71

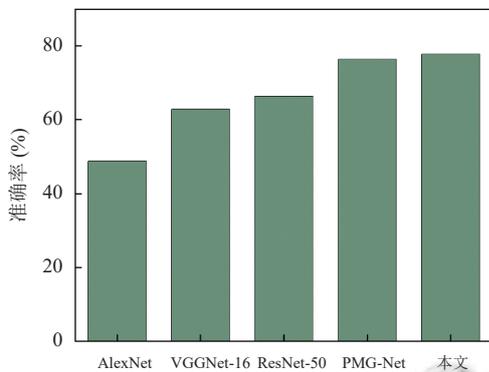


图 11 分类网络的正确率对比图

表 2 Transportation 数据集实验结果 (%)

方法	基准模型	准确率
PMG-Net	ResNet-50	76.34
	ResNet-101	77.01
本文网络	ResNet-50	76.92
	ResNet-101	77.71

实验的准确率对比图如图 13 所示. 与没有使用拼图生成器的结果相比, 本文使用的网络模型的结果仍然获得了更好的结果. 在 ResNet-50 网络中, 本文的方法在 Medical 数据集上准确率提高了 0.97%, 在 Transportation 数据集上提高了 0.75%. 在 ResNet-101 网络中, 在 Medical 数据集中本文方法优于没有使用拼图生

成器的方法 1.99%, 在 Transportation 数据集中性能优于其 0.55%. 这证明了拼图生成器可以提升本文模型的性能, 它的存在使模型的整体性能有了很明显的提高, 有很大的优势.

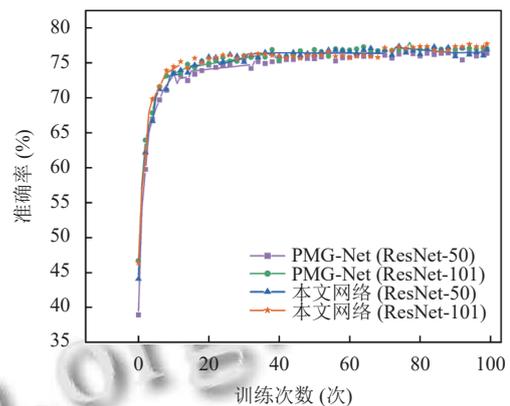


图 12 各网络的正确率对比图

表 3 准确率对比 (%)

准确率	网络模型	Medical 数据集	Transportation 数据集
Accuracy_a	改进的PMG-Net (ResNet-50)	69.65	76.17
	改进的PMG-Net (ResNet-101)	70.48	77.16
Accuracy	改进的PMG-Net (ResNet-50)	70.62	76.92
	改进的PMG-Net (ResNet-101)	72.47	77.71

#### 4 结论

本文提出利用 PMG-Net 网络解决 logo 图像分类的方法, 使用细粒度图像分类的方法解决 logo 图像分类的问题. 传统的卷积分类方法只把最后一层的输出作为指标进行图像的分类, 而本文的网络结构是一种

新颖的渐进式训练策略,它实现了将多粒度信息融合在一起以实现分类的准确性.多粒度的信息有助于避免类内较大变化的影响,真正做到了不仅是识别零件,还聚焦于每个部分真正有区别的部分,提高了网络模型分类的准确率.为了更好地辅助分类,本文采用了拼图生成器这一方法,在网络模型中加入了不同尺度的拼图块,生成包含不同粒度级别信息的输入图像,可以让网络模型更关注细节特征,从而更好地获得判别性区域的信息.实验表明,本文方法在 logo 图像分类数据集上的表现优于传统方法,为后续将 logo 信息用于版权保护、交通智能化、医疗信息智能化等方面,实现更精确的 logo 检测提供了新的思路.但本文提出模型在数据集上对不同种类图案依然存在分类效果和性能差距明显,在今后的研究中需要进一步改进模型,研究为不同属性类别设置合适的权重系数,解决分类效果差距明显的问题.

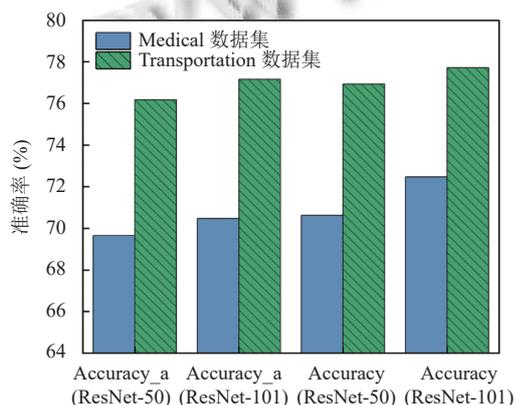


图 13 准确率对比图

### 参考文献

- 1 Revaud J, Douze M, Schmid C. Correlation-based burstiness for logo retrieval. Proceedings of the 20th ACM International Conference on Multimedia. Nara: ACM, 2012. 965–968.
- 2 Boia R, Florea C, Florea L, *et al.* Logo localization and recognition in natural images using homographic class graphs. Machine Vision and Applications, 2016, 27(2): 287–301. [doi: 10.1007/s00138-015-0741-7]
- 3 Kumar NV, Pratheek, Kantha VV, *et al.* Features fusion for classification of logos. Procedia Computer Science, 2016, 85:

- 370–379. [doi: 10.1016/j.procs.2016.05.245]
- 4 Bianco S, Buzzelli M, Mazzini D, *et al.* Deep learning for logo recognition. Neurocomputing, 2017, 245: 23–30. [doi: 10.1016/j.neucom.2017.03.051]
- 5 Hou SJ, Lin JW, Zhou SB, *et al.* Deep hierarchical representation from classifying logo-405. Complexity, 2017, 2017: 3169149. [doi: 10.1155/2017/3169149]
- 6 Du RY, Chang DL, Bhunia AK, *et al.* Fine-grained visual classification via progressive multi-granularity training of jigsaw patches. Proceedings of the 2020 European Conference on Computer Vision. Glasgow: Springer, 2020. 153–168. [doi: 10.1007/978-3-030-58565-5\_10]
- 7 Wei C, Xie LX, Ren XT, *et al.* Iterative reorganization with weak spatial constraints: Solving arbitrary jigsaw puzzles for unsupervised representation learning. Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019. 1910–1919.
- 8 Karras T, Aila T, Laine S, *et al.* Progressive growing of GANs for improved quality, stability, and variation. arXiv:1710.10196, 2017.
- 9 Ahn N, Kang B, Sohn KA. Image super-resolution via progressive cascading residual network. Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Salt Lake City: IEEE, 2018. 791–799.
- 10 Wang YF, Perazzi F, McWilliams B, *et al.* A fully progressive approach to single-image super-resolution. Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Salt Lake City: IEEE, 2018. 864–873.
- 11 Karras T, Laine S, Aila T. A style-based generator architecture for generative adversarial networks. Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019. 4396–4405.
- 12 Shaham TR, Dekel T, Michaeli T. SinGAN: Learning a generative model from a single natural image. Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision. Seoul: IEEE, 2019. 4569–4579.
- 13 Hu J, Shen L, Sun G. Squeeze-and-excitation networks. Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018. 7132–7141.

(校对责编: 孙君艳)