

基于改进 ResNet-18 的京剧旦角服饰分类研究

黄若曦

暨南大学管理学院, 广州 510632, 广东, 中国

摘要: 为建立京剧服饰数据库、保护戏曲服饰的文化遗产价值, 本文针对分类准确率不高的问题, 提出一种基于改进 ResNet-18 的京剧旦角服饰分类模型 (ResNet18-APSS)。首先, 收集京剧旦角服饰样本图像, 构建包含 18 个类别的数据集; 其次, 对预训练的 ResNet-18 模型进行改造, 以 APSS 模块替换其第四层卷积模块; 再次, 采用 Optuna 进行超参数寻优以获得最优配置, 并引入迁移学习与自动混合精度训练策略以进一步提升模型性能, 同时将损失函数调整为交叉熵损失函数以加快收敛速度; 最后, 使用寻优得到的最优超参数完成训练与验证。在自建京剧旦角服饰数据集上的对比结果表明, 改进后的模型能够有效实现京剧旦角服饰分类, 平均准确率达到 95.41%。本研究为京剧旦角服饰的分类难题提供了有效解决方案, 对推进其数字化进程具有重要意义。

关键词: 京剧; 京剧服饰; 服饰图像识别; ResNet-18; 迁移学习

Dan Character Peking Opera Costume Classification Based on Improved ResNet-18

RuoXi Huang

School of Management, Jinan University, Guangzhou 510632, Guangdong, China

Abstract: To establish a Peking opera costume database and preserve the cultural heritage of opera costumes, an improved ResNet-18-based classification model (ResNet18-APSS) for classifying Peking opera Dan costumes is proposed, addressing the issue of classification accuracy. First, Peking opera Dan costume sample images were collected to construct a dataset comprising 18 categories. Then, the pre-trained ResNet-18 model was modified by replacing the fourth convolutional layer with an APSS module. Hyperparameter optimization was performed using Optuna to find the optimal configuration. Transfer learning and automatic mixed precision training strategies were employed to enhance model performance further, and the loss function was adjusted to CrossEntropyLoss to accelerate convergence. Finally, the identified optimal hyperparameters were used for training and validation. Comparative results on the self-constructed Peking opera Dan costume dataset demonstrate that the improved model effectively classifies Peking opera Dan costumes, achieving an average accuracy of 95.41%. This study provides an effective solution to the classification challenges of Peking opera Dan costumes and significantly contributes to advancing their digitization.

Keywords: Peking Opera; Peking Opera costumes; Costume image recognition; ResNet-18; Transfer learning



一、引言

京剧服饰是中国戏曲的代表，构成表演艺术的重要组成部分。作为京剧的载体，这些服饰已逐渐发展为一门精致的艺术 [1]。中国传统戏曲服饰起源于明代，由明代生活服装经提炼、夸张与装饰演变而来，以满足叙事与表演的需要。京剧服饰研究可为中国戏曲服饰的传承与发展提供理论支撑。由于女性角色身份与性格的多样性，作为京剧四大行当之一的旦角进一步细分为六个更具体的类别——青衣、花旦、花衫、武旦、老旦与刀马旦。旦角服饰做工精美、色彩丰富、纹样华丽，是京剧服饰艺术的集中体现。然而对于普通观众而言，要在众多旦角服饰类型之间加以辨别仍具有相当难度。

依据穿戴规制，京剧服饰可分为蟒、帔、褶、靠、衣五大类 [2]。这些类别又依角色行当与人物身份进一步细分。在旦角服饰中，“蟒”可细分为女蟒、旗蟒、老旦蟒，见图 1；“帔”可细分为女帔、老旦帔；“褶”可细分为女褶子、老斗衣；“靠”可细分为女硬靠、女改良靠；“衣”可细分为袄裙、袄裤等。

随着信息技术的快速发展，数字化因其非破坏性与广泛适用的优势，已成为文化遗产保护与传播的新兴趋势。对于种类繁多的京剧旦角服饰而言，准确有效的分类模型是其数字化存储的基础。当前京剧服饰研究主要集中在创新设计、纹样数字化与增强现实交互等方面。



Figure 1 Physical Layouts of Mang

图 1 蟒的实物形制

在京剧服饰的自动分类方面，Liu 等基于改进的 YOLOv5 模型开展了分类研究，但该研究仅将戏曲服饰粗略划分为五大类，未能实现对旦角服饰的精细识别。

支持向量机（SVM）、K 近邻（KNN）与决策树算法等传统机器学习方法需要大量预处理工作，且分类准确率高度依赖特征选择的质量。随着深度学习的兴起，卷积神经网络（CNN）已成为服装分类任务的主流方法。CNN 能够从原始图像中自动学习高层特征表示，避免了人工特征提取的复杂性。Seo 和 Shin 率先在 CNN 模型中引入分层结构用于服装图像分类，提出 H-CNN 模型，显著提高了分类准确率并提升了整体分类性能。Su 提出一种将 Hu 不变矩与 CNN 特征相结合的服装分类算法，该方法采用信息增益进行特征选择，在多任务场景下的表现优于传统单任务模型。Obaid 和 Jasim 对 Fashion-MNIST 数据集进行了扩充与增强，构建了五层卷积深度神经网络，其性能较 VGG16 与 AlexNet 模型有显著提升。深度学习模型能够准确捕捉服装图像中的纹理、色彩、款式与纹样等复杂特征，从而显著提高任务准确率 [3]。Shin 等提出一种多层卷积与池化层优化的 CNN 模型，利用自适应动态学习率加速模型收敛，在 Fashion-MNIST 数据集上的实验表明，该方法显著提高了分类准确率并有效减轻了过拟合现象。

在传统服饰分类领域，也涌现出一些基于深度学习的传统服饰分类与检索方法。Nawaz 等提出一种基于 Google Inception 架构与 RmsProp 优化器的孟加拉传统服饰自动分类模型，该模型具有良好的泛化能力，能够在

图像质量不同的条件下实现传统服饰的分类。类似地, Athala 等采用一种新颖的数据增强策略, 通过对样本图像随机旋转与缩放以模拟不同角度和尺寸的图像, 从而增强分类的鲁棒性, 实现了对复杂印尼传统服饰纹样的自动识别与分类。Rangkuti 等在执行印尼传统服饰纹样自动分类任务时, 更加关注模型的可靠性与鲁棒性; 他们通过展示 InceptionV3 与 ResNet50 V2 在小样本数据集下的高准确率, 验证了 Inception 模型在传统服饰纹样分类中的鲁棒性。

本文提出一种基于 ResNet-18 的京剧旦角服饰自动分类方法, 实现了复杂背景下旦角服饰的高精度识别。在原有算法中集成新的模块, 提升了识别准确率, 解决了复杂环境中旦角服饰识别的难题。本研究在京剧旦角服饰分类方面的目标如下: (1) 建立京剧旦角服饰标注体系, 构建完备的数据库; (2) 提出更为有效的旦角服饰识别方法, 帮助观众更好地理解这些服饰; (3) 为京剧艺术的传承与传播作出贡献。

二、研究方法

改进后的 ResNet-18 网络结构 (ResNet18-APSS) 以原始 ResNet-18 的架构为基础。ResNet-18 包含五个阶段: 初始卷积与池化层、四个残差模块 (Residual Block) 以及一个全连接层。ResNet-18 的核心结构是残差块, 每个残差块包含两个卷积层, 每个卷积层之后依次为批归一化与 ReLU 激活函数。

在 ResNet-18 中, 每一阶段的第一个残差块为下采样残差块, 第二个为标准残差块。残差块的输出可表示为:

$$y = F(x) + x \quad (1)$$

式中, $F(x)$ 表示卷积层与激活函数的组合, x 表示输入图像。

在这一设计下, 即使网络未能学到所期望的特征, 仍可通过直接传递 x 而避免性能退化。该结构使网络更易于优化, 并使更深层网络的高效训练成为可能。

与传统卷积神经网络相比, 残差连接能够缓解深层神经网络中的梯度消失或梯度爆炸问题, 从而解决深层网络的退化问题。通过堆叠多个浅层块形成深层网络, 可以更好地捕捉图像中的复杂特征, 进而提高分类准确率。

(一) ASPP 模块

空洞空间金字塔池化 (ASPP) 是一种常用的卷积神经网络模块, 多用于在图像分类任务中增强特征提取模块的多样性与鲁棒性。ASPP 模块通过将不同采样率的空洞卷积与图像级特征相结合, 增强了网络捕捉多尺度上下文信息的能力。对于不同尺寸的目标, ASPP 能够捕捉从局部到全局的上下文信息, 提升模型对多尺度目标的鲁棒性。

ResNet18-APSS 移除了原有的第四层卷积模块并以 ASPP 替代, 以增强特征提取的多尺度能力; 通过感受野的扩展与全局池化, 该模型能够更好地处理背景复杂的场景以及远近景物物体尺度差异较大的图像。

在 ResNet18-APSS 中, 采用交叉熵损失函数作为损失函数, 其表达式定义如下:

$$\text{CrossEntropyLoss} = -(1/N) \sum_{(i=1..N)} \sum_{(c=1..C)} y(i,c) \cdot \log(p(i,c)) \quad (2)$$

式中, N 表示该批次的样本数量, C 表示类别总数, $y(i,c)$ 表示二值指示变量 (取值 1 或 0), 用以标示类别 c 是否为样本 i 的正确类别; $p(i,c)$ 表示由 softmax 函数得到的样本 i 属于类别 c 的预测概率。该损失函数更适用于评价模型在分类任务中的分类性能; 此外, 它能够引导模型更有效地学习不平衡的类别分布, 在与 APSS 模块结合时可实现分类任务的更优优化。

(二) AdamW 优化器

ResNet18-APSS 采用 AdamW 作为优化器（AdamW 是在 Adam 优化算法基础上的改进）。通过计算梯度的一阶动量与二阶动量，自适应地调整各参数的学习率，从而提升模型的泛化性能。显式的权重衰减能够有效控制参数规模，进而降低过拟合风险。

梯度的一阶动量（ m_t ）：

$$m(t) = \beta_1 \cdot m(t-1) + (1 - \beta_1) \cdot g(t) \quad (3)$$

式中， m_t 表示梯度的指数加权平均。

梯度的二阶动量（ v_t ）：

$$v(t) = \beta_2 \cdot v(t-1) + (1 - \beta_2) \cdot g(t)^2 \quad (4)$$

式中， v_t 表示梯度平方的指数加权平均。

与 Adam 不同，AdamW 将权重衰减从梯度更新中分离出来，直接作用于参数：

$$\theta(t+1) = \theta(t) - \alpha \cdot m(t) / (\sqrt{v(t)} + \epsilon) - \alpha \cdot \lambda \cdot \theta(t) \quad (5)$$

式中， α 表示学习率，由此可避免 Adam 自适应优化对权重衰减的干扰 [4]。

（三）Optuna 超参数优化

超参数调优是传统机器学习项目中最为耗时的任务之一，其选择直接影响模型性能 [5]。鉴于京剧旦角服饰分类任务中网络结构复杂且需调优的超参数数量众多，本文将 Optuna 模块集成到 ResNet18-APSS 中，并对参数优化单元进行了重新设计，以最大化模型的分类准确率与整体性能。

Optuna 的核心是“目标函数”，其通常返回一个待优化的性能指标。每次运行试验时，以一组超参数调用“目标函数”，执行训练并返回结果。“目标函数”的每一次独立执行即为一次试验（trial），多次试验可在多个工作节点上并发运行。

Optuna 的建议算法通常采用随机搜索或 TPE（树结构 Parzen 估计器）算法。随机搜索通过下式寻找最优参数组合：

$$x(\text{best}) = \arg \min f(x(i)), \quad x(i) \in X \quad (6)$$

式中， $f()$ 为超参数的模型性能评价函数。

（四）OneCycle LR 学习率策略

OneCycle LR 策略是一种学习率调度策略，通过在训练过程中设置单周期的学习率变化，使模型收敛更快并降低过拟合风险 [6]。OneCycle LR 的超参数设置主要为 `max_lr` 与 `total_steps`，这减少了 Optuna 搜索超参数所需的时间。

（五）迁移学习

迁移学习是将已有模型的知识迁移到新任务上的方法，尤其适用于数据量小但特征复杂的新任务。在深度学习中，迁移学习通常先在大型数据集（如 ImageNet）上预训练深度神经网络，再将预训练模型应用于特定的目标任务，并以目标数据集进行微调，从而快速获得高性能的分类模型。

在 ImageNet 上预训练的 ResNet-18 模型已具备较强的特征提取能力，将其迁移至京剧旦角服饰分类任务时，可显著提高分类准确率并缩短训练时间。在实现过程中，本文采用微调策略优化 ResNet18-APSS 模型，使其能够适应京剧旦角服饰的独特特征。

三、实验

（一）数据增强与预处理

京剧旦角服饰数据集是评价与检验京剧旦角服饰分类性能的重要资源。然而，目前尚缺乏专门面向京剧旦角服饰图像的公开数据集。本文通过文献调研与网络检索，对蟒、帔、褶、靠等传统类别作了进一步细分，最终确定 18 种具体的京剧旦角服饰类型作为研究对象：女蟒、旗蟒、老旦蟒、女帔、老旦帔、女褶子、老斗衣、袄裙、袄裤、宫装、古装衣、女硬靠、女改良靠、水田衣、战袄战裙、鱼鳞甲、女罪衣与旗袍。研究从各类互联网平台收集了 1700 张京剧旦角服饰图像，以 JPG 格式保存，清晰度与质量均较高（图 2）。



Figure 2 Sample images

图 2 样本图像

为使结果更加稳健，研究选取部分样本进行人工标注，并按 80%与 20%的比例将数据集划分为训练集与测试集。

研究采用 Python-OpenCV 库对图像样本进行增强，包括调整图像亮度、饱和度与对比度，添加椒盐噪声，水平翻转等操作，将数据集扩充至 8250 张 [7]。为确保模型在测试时能够正常处理输入数据，对每个图像通道进行归一化，并分别采用基于 ImageNet 数据集计算得到的均值与标准差。为使测试时模型输入数据的形状与训练时保持一致，测试集图像同样被调整为 256×256 的固定尺寸。最终得到符合 ResNet-18 网络输入要求的样本图像，并以京剧旦角服饰数据集的形式加以存储 [8]。

（二）模型训练

实验所用硬件为 Intel AMD EPYC 9354 处理器（16 核）与 NVIDIA RTX4090 24GB GPU 显卡，采用 Python 语言、PyTorch 深度学习框架与 JupyterLab 开发平台。在仿真训练中，使用 Optuna 自动搜索学习率与批量大小，训练迭代轮数设为 10 次，搜索轮数设为 15 次；采用 AdamW 作为优化器，通过计算梯度的一阶矩与二阶矩调整各参数的学习率；损失函数选用交叉熵损失函数，训练迭代轮数设为 60 次，并采用 AMP 自动混合精度训练以加速模型收敛。评价指标采用准确率（ACC）、F1 分数与混淆矩阵。

四、结果与讨论

为验证数据增强的效果，在控制网络模型等其他因素的条件下，将经数据增强扩充后的数据集与未经数据增强的数据集分别送入同一 ResNet-18 网络进行训练。原始数据集模型的准确率为 88.83%，增强后数据集的准

确率为 90.83%，这表明数据增强有助于深度学习模型更好地学习京剧旦角服饰的特征，提升模型的识别效果。

Table 1 Ablation Study

表 1 消融实验

模型	ASPP	AdamW	OneCycle LR	ACC/%	F1 分数
ResNet-18	×	×	×	90.83	0.914
改进模型 1	√	×	×	91.76	0.919
改进模型 2	×	√	×	93.69	0.938
改进模型 3	×	×	√	93.40	0.942
ResNet18-APSS	√	√	√	95.41	0.958

为验证 OneCycle LR 与 AdamW 对迁移后 ResNet-18 模型的改进效果,本文基于数据增强后的数据集与迁移后的 ResNet-18 模型开展了三组对比实验,涉及 ResNet-18 模型、改进模型 1 (ResNet-18+OneCycle LR)、改进模型 2 (ResNet-18 模型+AdamW 优化器)、改进模型 3 (ResNet-18+APSS) 以及 ResNet18-APSS, 见表 1。首先,在原始 ResNet-18 网络中引入 OneCycle LR 学习率调度策略,模型分类准确率由 90.83%提升至 93.40%, F1 分数亦由 0.914 提升至 0.942。这些结果表明,通过动态调整学习率能够有效加快模型的收敛速度并提升模型性能。其次,采用 AdamW 优化器进一步提升模型性能,准确率达到 93.69%, F1 分数为 0.938,较初始的 SGD 优化器有所提高。最后,通过结合 Optuna 超参数优化、数据增强与自动混合精度训练,本文提出的改进模型取得了最佳的分类准确率与 F1 分数,平均分类准确率为 95.41%, F1 分数为 0.958。与原始 ResNet-18 模型相比,本研究通过多种改进措施显著提升了模型的性能与整体稳定性,见图 3。

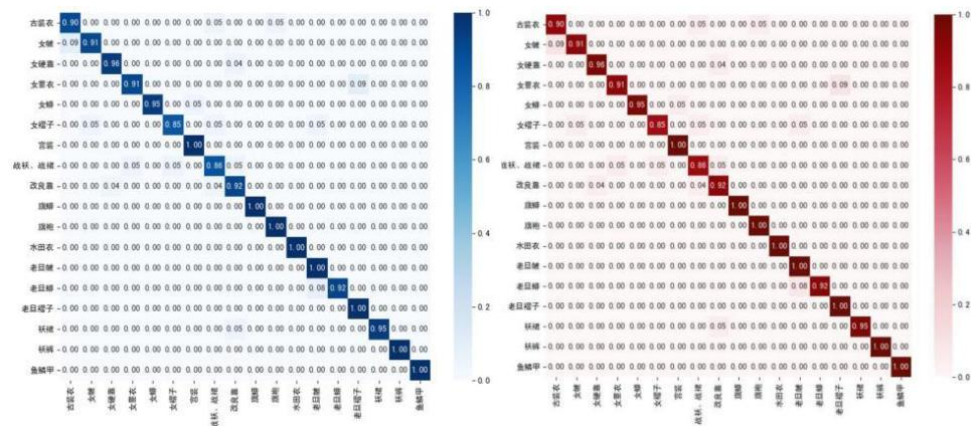


Figure 3 Confusion Matrix ((a) pre-trained ResNet-18; (b) ResNet18-APSS)

图 3 混淆矩阵 ((a) 预训练 ResNet-18; (b) ResNet18-APSS)

为全面评价模型性能, 本文将 ResNet18-APSS 与 VGG16、MobileNetV2、ResNet-50 以及 ShuffleNetV2 进行了对比。

ResNet18-APSS 融合了数据增强、超参数优化与自动混合精度训练, 其分类准确率显著优于上述模型。这表明改进后的模型在特征提取、鲁棒性与泛化能力方面具有明显优势, 在处理纹理与色彩复杂的京剧旦角服饰图像时尤为突出, 展现出更好的分类效果与稳定性。

五、结论

为解决京剧旦角服饰识别与分类效果不佳的问题,本文对 ResNet18 模型进行了改进,并将自建的京剧旦角服饰数据集送入改进后的模型进行训练,得到如下结论:

(1) 数据增强操作有效提升了改进模型的泛化能力,准确率提高了 2.00%。ASPP 模块的引入、学习率调整策略 OneCycle LR 的调整以及 AdamW 优化器的采用,均有利于提升模型的识别效果与准确率,分别较未改进的 ResNet-18 模型提高 0.93%、2.86%与 2.57%。通过迁移学习方法,将从 ImageNet 数据集中习得的知识迁移至京剧旦角服饰数据集,增强了模型对不同类型京剧旦角服饰的识别适应性。

(2) 与 ResNet-18 模型相比,ResNet18-APSS 的准确率提升了 4.58%,F1 分数提高了 0.040,实现了精确率与召回率的协调统一。该模型在京剧旦角服饰自动分类中表现优异,即使在数据有限的条件下仍能保持较高的准确率与稳定性,在传统服饰分类研究中显示出切实的应用价值。

(3) 建立京剧旦角服饰识别模型及京剧旦角服饰数据库中相应的十八类分类标签,能够加快京剧旦角服饰数字化资源的建设,对京剧旦角服饰文化的传承与发展产生积极影响。

然而,由于纹样、款式与工艺的差异,在这十八个类别之外仍存在更为细致的分类。考虑到舞台表演的复杂性与京剧的持续演变,对京剧旦角服饰的认识仍有进一步深化的空间。

未来关于京剧旦角服饰识别与分类的研究,将着重于提升数据质量、优化模型以及探索多场景应用,同时强调传统文化保护与技术创新的融合。

参考文献

- [1] 杨宜勇. 就业优先战略实施保障机制. 宏观经济管理, 2022(04): 28-35.
- [2] He K, Zhang X, Ren S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1904-1916.
- [3] 陈志武. 金融赋能实体经济边界研究. 金融研究, 2022(03): 1-18.
- [4] Lee Y S. A study of stage costume of Peking Opera. The International Journal of Costume Culture, 2003, 6(1): 38-51.
- [5] Li J. Research on digital display of intangible cultural heritage. Journal of Education, Humanities and Social Sciences, 2023, 8: 54-61.
- [6] Li L, Jie Y, Yanxuan J, et al. Clothing image classification algorithm based on improved MobileNet v2. Advanced Textile Technology, 2024, 32(4): 93.
- [7] Lian X, Pang Y, Han J, et al. Cascaded hierarchical atrous spatial pyramid pooling module for semantic segmentation. Pattern Recognition, 2021, 110: 107622.
- [8] 田卫平. 区域文化产业差异化发展路径. 艺术百家, 2022, 38(02): 145-151.