

文章编号: 2095-2163(2022)02-0126-07

中图分类号: TP 391.1

文献标志码: A

基于深度特征和注意力机制的艺术图像情感分类研究

杨松^{1,2,3}, 刘佳欣¹, 吴桐¹

(1 大连外国语大学 软件学院, 辽宁 大连 116044; 2 大连外国语大学 语言智能中心, 辽宁 大连 116044;

3 网络空间大数据安全院级研究中心, 辽宁 大连 116044)

摘要: 作为人类精神活动产物的艺术图像,其本身蕴含着丰富的情感语义信息,研究艺术图像的情感分类有助于艺术图像的鉴赏与保护,以图像为对象的情感分类研究已成为情感计算的研究热点,但该分类主要依赖于图像低层特征的抽取,从而导致图像情感分类结果不高。本文提出了一种基于底层特征和注意力机制的艺术图像情感分类模型,即提取艺术图像的CLAHE颜色特征、Laplacian纹理特征与艺术图像深度特征融合作为输入,同时引入CBAM注意力机制以关注图像的重点区域,构建面向艺术图像情感分类的卷积神经网络模型FeaturesNet。实验结果表明,应用本文提出的模型在艺术图像数据集上的情感分类准确率可达93.36%,相较于其他模型有较大提升。

关键词: 艺术图像; 情感分类; 卷积神经网络; 深度特征; 注意力机制

Research on emotion classification of artistic images based on deep features and attention mechanism

YANG Song^{1,2,3}, LIU Jiaxin¹, WU Tong³

(1 School of Software, Dalian University of Foreign Languages, Dalian Liaoning 116044, China; 2 Research Center for Language Intelligence, Dalian University of Foreign Languages, Dalian Liaoning 116044, China; 3 Research Center for Networks Space Multi-Languages Big Data Intelligence Analysis, Dalian University of Foreign Languages, Dalian 116044, China)

[Abstract] Emotion classification based on image has become a research hotspot of emotion computing. However, the classification of image emotion mainly depends on the extraction of deep features, which leads to low classification results. As the product of human spiritual activities, art images itself contains rich emotional semantic information. Research on the emotional classification of art images is helpful to the appreciation and protection of art images. An emotion classification model of art images based on deep-level features and attention mechanism network is proposed in this paper which extracts the CLAHE color features and Laplacian texture features and fuses them into their deep-level features of art images as inputs. At the same time, an attention mechanism based on CBAM is introduced to pay close attention to important areas, and a model named as FeaturesNet is constructed for classifying emotion of art images. The experimental results show that the emotion classification accuracy of this improved convolutional neural network model on the image data set of art images defined in this paper can reach 93.36%, which is greatly improved compared with other models.

[Key words] artistic images; emotional classification; convolutional neural network; deep features; attention mechanism

0 引言

随着微博、微信、论坛等社交媒体的广泛使用,越来越多的用户喜欢以文本、图像及视频的形式分享信息、表达观点和宣泄情绪。网络中的海量信息往往存在很多主观的感情色彩,基于文本的情感分析技术已在信息情感的把握中发挥了重要作用,而基于图像的情感分析技术仍有待深入的研究。图像情感分类作为计算机视觉、模式识别、认知科学等多个学科的交叉技术,近年来受到国内外研究人员的广泛关注^[1]。希望借助计算机技术研究和建立符

合人类感知的情感映射机制,理解并准确分类图像中蕴含的高层情感语义^[2]。目前在一些特定场景的图像情感分类研究已取得一些成果,例如人脸表情、自然风景、服饰面料等,而以艺术图像为目标的情感分类研究相对较少。艺术图像作为艺术作品的重要组成部分,本身具有丰富的情感语义信息,能够通过颜色搭配、线条手法、纹理特征及抽象主体特写等方式向人们传递作品内在的情感,相对于文本情感,图像情感表现得更为抽象、理解的难度更大。研究艺术图像的情感识别与分类,对艺术图像的鉴赏、保护等具有重要指导意义。

基金项目: 国家自然科学基金(61806038);辽宁省社会科学规划基金(L18BTQ005);辽宁省教育厅科学研究项目(2019JYT07)。

作者简介: 杨松(1977-),男,博士,副教授,CCF会员,主要研究方向:图像情感计算、模式识别;刘佳欣(1996-),女,硕士研究生,主要研究方向:图像情感识别;吴桐(2001-),女,硕士研究生,主要研究方向:图像情感识别。

收稿日期: 2021-10-04

本文从情感角度出发,以艺术图像作为研究对象,通过深度学习来挖掘艺术作品带给欣赏者的情感变化,提出一种基于深层特征和注意力机制的艺术图像情感分类模型 FeaturesNet,抽取艺术图像的 CLAHE 颜色特征和 Laplacian 纹理特征,与学习到的深度情感语义特征融合,弥补情感语义特征抽取的不足。引入 CBAM 注意力机制,在降低模型过拟合的同时提高对艺术图像情感分类的准确率,为基于深度学习的图像情感分类提供一定参考。

1 相关研究综述

近年来,国内外的一些学者在艺术图像情感分类领域取得了一些成果。雷晶晶^[3]以木板套色版画为研究对象,提出基于 SVM 算法和模糊聚类算法的图像情感分类模型,实验采用人工标注的数据集,由于数据集小导致情感分类的准确率不高;李博等^[4]提取颜色直方图、Itten 对比、二阶矩特征,验证加权 K 近邻算法在 Abstract 数据集上离散情感分布预测的有效性,但参数 K 需要人工选取。传统的机器学习算法在图像情感语义分类中的应用已达到瓶颈,深度学习模型被引入到图像情感语义分类领域,能够解决图像特征需要人为构造的问题,有利于缩小图像底层视觉特征和深层语义特征之间的“语义鸿沟”^[5]。

盛家川等^[6]将人类认知与 CNN 模型结合,根据中国画显著性区域和笔道复杂度提取感兴趣区域,利用预训练的微调 GoogLeNet 模型提取中国画的深层特征并进行可视化,结合中国画情感表达手法知识,调整模型结构,优化训练参数,提升中国画情感分类的准确率;盛家川等^[7]提出基于 CNN 特征聚合重标定的中国画情感分类算法,重定向图像为正方形,在 ResNet 模型上融合同一模块下各卷积层特征并反馈至卷积层输出,重新标定通道特征值并强化不同特征的激活强度。将预测值作为权重,计算各类别激活映射,加权融合所有类别的激活定位,视觉解释情感区域特征,实现对中国画的情感特征可视化及情感分类任务。白茹意等^[8]利用抽象画的基本元素与人类情感之间存在的关系,提取抽象画的底层视觉特征,采用迁移学习方法,将预训练模型的参数迁移至目标模型并对模型进行微调,提取抽象画的深层语义特征,最后将两者线性融合,利用 SVM 方法实现抽象画的情感分类任务;王征等^[9]利用 AlexNet 模型提取中国画的深层特征并与颜色特征融合,将得到的特征组合进行特征处理,使用

SVM 分类器进行情感分类;张浩等^[10]为深入分析云南少数民族绘画作品所蕴含的情感,微调预训练 VGG16 模型,将该模型应用于 Twitter 数据集中训练,以提升模型学习能力,采用不同方法分别对绘画作品进行数据增强,输入训练好的微调模型中,比较各方法以及组合处理方法对民族绘画作品的情感分类结果,以达到情感分类的最优状态。

深度卷积神经网络属于有监督学习方法,分类效果依赖于大量样本数据的训练,通过模型训练提取图像深度特征并与传统的图像底层特征融合,可以给模型性能带来一定的改善。同时,图像情感的极性仅与图像特定的区域有关,而其他区域对分类的影响极小,通过在深度学习中引入视觉注意力机制,使模型更加关注那些影响图像情感分类结果的关键区域。此外,随着模型层数逐渐加深,训练过程容易造成图像细节信息的丢失,改进模型结构仍存在着巨大的提升空间。图像情感分类可以在图像深度特征融合、注意力机制引入、模型结构改进等方面开展深入研究,提升图像情感分类的准确率和效率,同时降低模型的过拟合度。

2 艺术图像的特征提取

2.1 CLAHE 颜色特征

颜色特征是图像最基本的底层视觉特征,是图像中关键与敏感的视觉信息。在心理学的相关研究中,有学者认为颜色是表达情感的普遍方式,具有艺术性、鲜明性、联想性、独特性、合适性及唤醒性等特点。积极情感的图像整体色彩较为鲜艳明亮,消极情感的图像颜色则暗淡压抑。在研究图像底层颜色特征时,大部分学者习惯采用颜色直方图表达颜色特征,可直观看出图像像素的分布情况。在同一数据集中,图像的像素分布存在较大差异,针对像素分布不均匀的图像,部分区域与整体图像相比较为明亮或光线较暗,因此采用普通颜色直方图均衡进行全局图像增强,效果不太理想。自适应直方图均衡(AHE)与颜色直方图相比,优势在于该方法通过计算图像每一处显著区域的直方图,重新调节图像亮度值,进行均匀分布,以改善图像每个区域的局部对比度,增强图像边缘的清晰度,但该方法容易导致图像产生噪声,进而产生图像过度增强的现象。限制对比度自适应直方图均衡(CLAHE)与 AHE 相比,通过对每一像素邻域的对数限制,得到对应的变换函数,可有效抑制图像噪声。与原图像相比,经 CLAHE 方法处理后的图像细节更为突出,特征更加丰富,颜色直方图较原图像相比也

更为平滑,如图1所示。

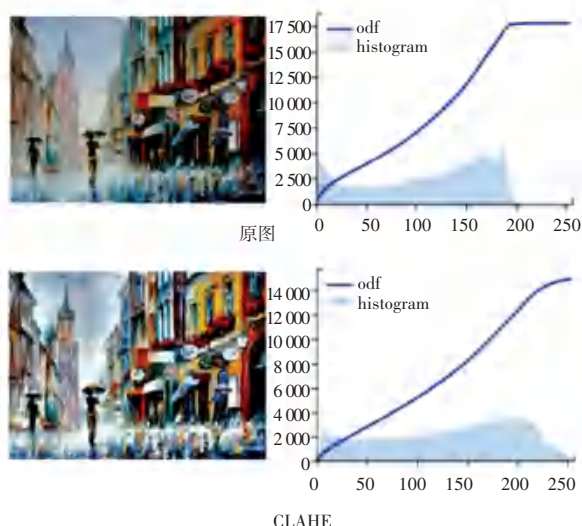


图1 CLAHE颜色特征可视化图

Fig. 1 Figure of CLAHE color feature visualization

经过 CLAHE 算法处理后的颜色特征输入到两个卷积层、两个最大池化层、一个全连接层、3 个 BN 层和 3 个激活层的网络中。其中,卷积核大小为 3×3 , $filters = 64$, $stride = 2$, $padding = 1$; 最大池化层均为 2×2 , $stride = 2$; 全连接层神经元为 1 000 个; 激活层使用 *Relu* 激活函数; 卷积层和全连接层后均连接 BN 层, 用于防止模型过拟合。将该网络模型命名为 ColorNet, 输出结果即为图像的颜色特征, 记为 X_1 , CLAHE 颜色特征提取过程, 如图 2 所示。

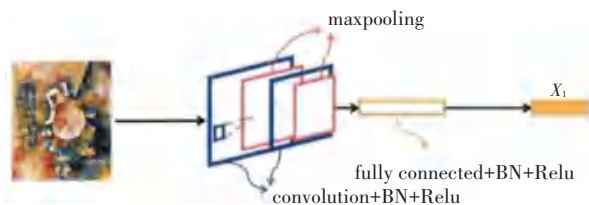


图2 CLAHE颜色特征提取过程

Fig. 2 Extraction process of CLAHE color features

2.2 Laplacian 纹理特征

通常物体的表面都具有纹理特征, 包含了对比度和空间频率等因素, 其目的是描述图像相邻像素之间灰度空间的遍历情况。图像纹理特征同属于图像底层特征, 虽然视觉效果没有颜色特征更直观, 但不同的纹理特征同样能够对人类的情感产生影响。拉普拉斯算子 (Laplacian) 是图像邻域内像素灰度差分计算的基础, 利用二阶微分运算推导出的一种图像邻域增强算法, 可用于描述图像区域的变化速度。其基本思想是当邻域的中心像素灰度低于其在邻域内的其他像素的平均灰度时, 中心像素的灰度应该进一步降低; 反之进一步提高中心像素的

灰度, 进而实现图像锐化处理^[11]。与一阶微分运算相比, 拉普拉斯算子边缘定位能力更强, 获取的边界信息更加细致。因此, 利用拉普拉斯算子对图像进行锐化边缘检测、提取图像纹理特征有助于图像分类、目标检测识别、图像分割等任务^[12]。二维图像拉普拉斯算子公式(1):

$$\nabla^2 f(x, y) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \quad (1)$$

其离散化形式可表示为式(2):

$$\nabla^2 f(x, y) = 4f(x, y) - f(x+1, y) - f(x-1, y) - f(x, y-1) - f(x, y+1) \quad (2)$$

Laplacian 算子对噪声很敏感, 为了获取较好的纹理, 突出边缘特征信息, 使用高斯滤波对图像进行降噪处理。本文提取图像纹理特征具体步骤为:

- (1) 利用高斯滤波去除图像噪声;
- (2) 将降噪后的图像转为灰度图像;
- (3) 使用 Laplacian 算子提取处理后的灰度图像的边缘特征;
- (4) 将处理后的灰度图像与边缘特征图像融合。

经 Laplacian 算子等处理后得到的纹理特征输入两个卷积层、两个最大池化层、一个全连接层、3 个 BN 层和 3 个激活层的网络中。其中, 卷积核为 3×3 , $filters = 64$, $stride = 2$, $padding = 1$; 最大池化层均为 2×2 , $stride = 2$; 全连接层神经元为 1 000 个; 3 个激活层使用 *Relu* 函数; 卷积层和全连接层后均连接 BN 层, 防止过拟合。将该网络模型命名为 LapNet, 输出结果即为图像的纹理特征, 记为 X_2 , Laplacian 纹理特征的提取过程, 如图 3 所示。

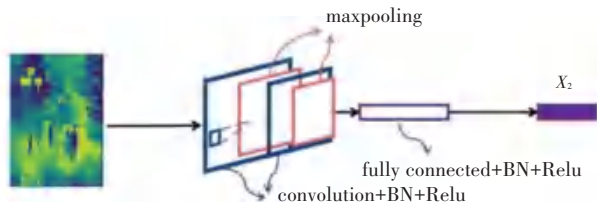


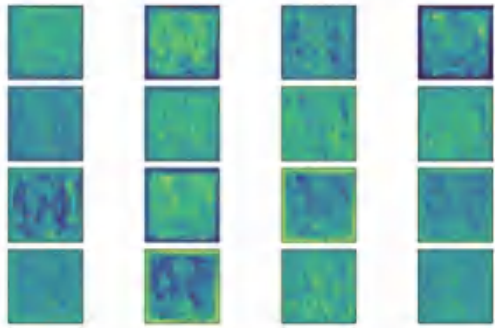
图3 Laplacian纹理特征提取过程

Fig. 3 Extraction process of texture features

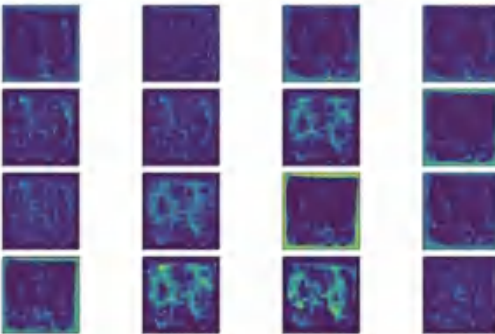
2.3 深度语义特征

在图像深度语义特征提取的过程中, 采用 ResNet101 与 CBAM 注意力机制融合的网络模型对数据集进行训练, 在不改变 ResNet101 模型原结构的前提下, 将通道注意力机制和空间注意力机制融入 ResNet101 模型的第一层卷积层和最后一层卷积层中, 集中关注图像主体区域和局部区域的特征信息, 充分挖掘艺术图像的整体特征和细节特征, 使模型训练和测试的准确率较传统卷积神经网络模型有

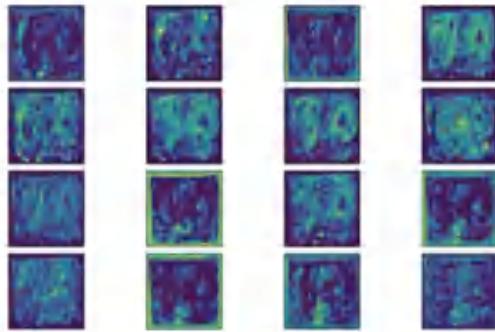
所提升。将该融合模型所提取的图像深层情感语义特征记为 X_3 , 在 ResNet101 与 CBAM 注意力机制融合模型中输入特征图后, 模型不同层所提取到的特征可视化效果图如图 4 所示。由此可见, 随着融合网络模型层数的不断加深, 肉眼可见的可视化效果逐渐模糊, 图像深层特征变得更加抽象。



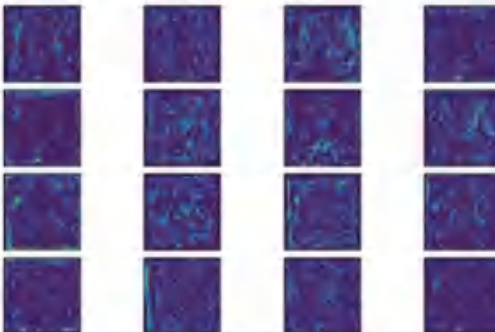
(a) Conv1层可视化图



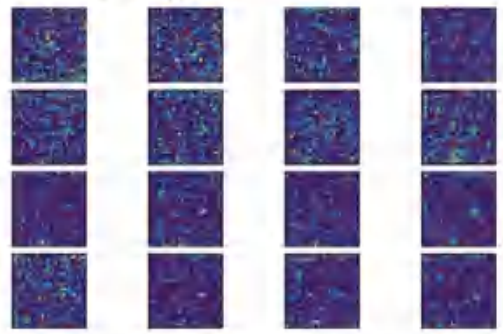
(b) Relu层可视化图



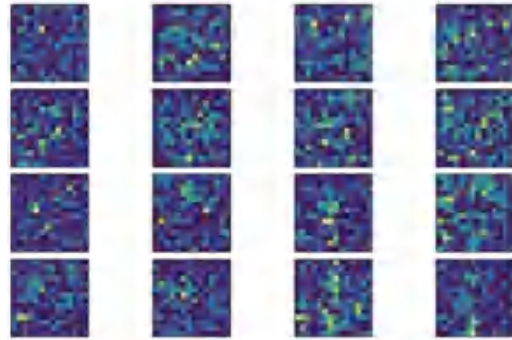
(c) MaxPool层可视化图



(d) Layer1层可视化图



(e) Layer2层可视化图



(f) Layer3层可视化图

图 4 不同层深层特征可视化图

Fig. 4 Visualization map of deep features in different layers

2.4 特征融合输入

将上述提取到的 CLAHE 颜色特征 X_1 、Laplacian 纹理特征 X_2 和 CBAM 深层语义特征 X_3 利用 `torch.cat()` 函数并按照维数 1(列) 进行拼接融合, 整体输入到含有两个全连接层、两个 BN 层、一个激活层的网络中, 将该卷积神经网络模型命名为 FeaturesNet, 其结构如图 5 所示。

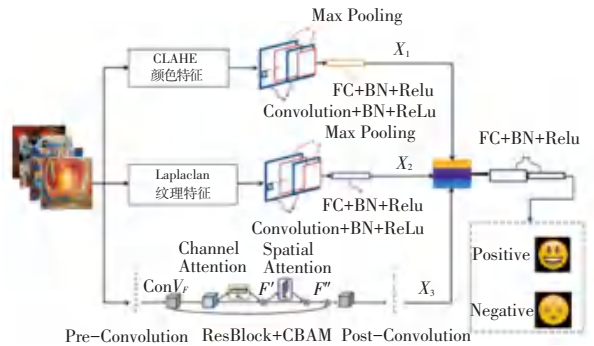


图 5 FeaturesNet 网络模型结构

Fig. 5 Network model structure of FeaturesNet

3 情感分类模型

3.1 ResNet

2015 年, 微软团队提出了 ResNet 模型, 模型结构较其他 CNN 模型更深, 整体性能也有所提高。该模型的基本残差模块由短路连接路径和残差路径组

成,通过引入网络前层与网络后层的短路连接路径加深模型的网络深度,有助于训练过程中梯度的反向传播,解决其他模型容易引发的梯度消失问题,提升深层次模型的网络性能。残差路径结构包括卷积层、BN层和ReLU激活层,在连续两层卷积层中,将输入特征 x 与卷积层输出特征合并,导致输出结果由原来两层输出 $F(x)$ 变成 $F(x) + x$ 。基于短路连接路径部分没有引进其他参数增加计算复杂度,因此即使增加模型层数,也能加快模型的训练速度,提升模型的训练效果。ResNet模型共有5种网络结构,其中ResNet101为残差神经网络中最常见的模型,其短路连接路径中共有3个卷积层,分别为两层 1×1 卷积层和一层 3×3 卷积层,通过这3层卷积层之间的短路连接路径加深各网络结构。

3.2 注意力机制

在计算机视觉领域中,注意力机制(Attention)与人类视觉机制十分相似,将关注焦点聚焦在图像局部显著区域位置。早期研究中,L Itti等^[13]受灵长类生物的视觉系统启发,提出视觉注意力机制模型,构建并融合颜色、方向和亮度特征图,最终提取复杂场景图像的显著区域位置。随着深度学习的普及,将注意力机制与卷积神经网络结合,通过快速扫描数据集,判定主要特征区域,并将注意力集中分布在感兴趣区域或重要信息特征位置,忽略非感兴趣区域及无关信息,可更快、更准确地获取数据集的主要特征信息,提升网络模型工作效率。CBAM注意力机制是由Sanghyun Woo等人于2018年提出的,其本质是前馈卷积神经网络,由通道注意力机制(ChannelAttention)和空间注意力机制(SpatialAttention)混合组成^[14]。其工作原理是首先向模型中输入一张特征图,CBAM注意力机制依次通过通道和空间两个维度提取特征图对应的注意力特征图;将注意力特征图与输入特征相乘得到自适应优化特征^[14]。在通道注意力机制和空间注意力机制模块内均包含池化操作,利用平均值池化和最大值池化对图像全局区域和局部区域给予关注,同时保留特征图细节处的语义信息。考虑到CBAM注意力机制具有轻量性和通用性的特点,本文将CBAM注意力机制与ResNet101卷积神经网络模型结合,通过端到端的训练,提升传统卷积神经网络模型的性能。

3.3 算法流程

将实验整体网络模型命名为FeaturesNet,输入的图像大小为 $224 \times 224 \times 3$ 。由于处理CLAHE颜色

特征和Laplacian纹理特征均在OpenCV库中进行,因此要将Tensor向量转换为numpy数组形式,再分别进行底层特征处理;将CLAHE颜色特征和Laplacian纹理特征再转回Tensor向量形式,分别输入ColorNet和LapNet网络中,进一步提取图像底层特征,ColorNet输出的特征记为 X_1 ,LapNet输出的特征记为 X_2 ,图像数据输入ResNet101与CBAM融合网络模型中得到的深层语义特征记为 X_3 ;最后,将得到的图像底层特征和深层情感语义特征融合,输入ClassNet网络中,进而对艺术图像进行情感语义分类。

4 实验结果与分析

4.1 实验环境

本文实验环境:处理器为Intel i5-8279U,主频2.40 GHz,实验基于Python 3.7的深度学习框架Pytorch,开发工具为PyCharm,使用Linux操作系统,GPU进行训练,GPU为NVIDIA Tesla V100-SXM2-16 GB,显存总量为16 160 MiB。

4.2 实验数据集的构建

缺少图像情感数据会导致卷积神经网络模型训练的不充分或模型易出现过拟合的问题。实验使用Machajdik等在2010年所创建ArtPhoto和Abstract两个公开数据集,其中ArtPhoto数据集由806幅艺术图像及情感标签组成,Abstract数据集包含280张艺术抽象画。由于ArtPhoto和Abstract数据集情感标签完全相同,积极类情感和消极类情感数据集划分方式一致,通过对积极类情感和消极类情感的标签进行投票统计,将得票数高的标签作为图像的最终情感类别,若积极类情感和消极类情感求和票数相同,则剔除该图像。经情感标签重新划分后,最终将ArtPhoto数据集和Abstract数据集合并作为本文实验的数据集,其中Positive类共540张图像,Negative类共538张图像。数据集各标签的示例图像如图6所示。

由于原数据集大多数图像格式是通道数为3的RGB格式图像,其中也存在少数通道数为1的灰度图像,为方便后续实验的展开,在数据预处理阶段统一将图像转为RGB格式。

情感语义信息的主观性和复杂性导致带有情感语义标签的图像数据集较少,而利用深度学习方法训练模型的本质是在海量数据集中学习深层次图像特征,进而获取图像情感语义信息,因此在数据集预处理阶段,还需对其进行数据增强操作。具体操作

包括对数据集中各图像翻转及旋转操作,数据集的增强示例图像如图 7 所示。对数据增强后的数据集进行人工筛选,最终形成各情感标签均为 1 920 张实验数据集。



图 6 图像情感数据集示例图像

Fig. 6 The schematic figures of image sentiment dataset

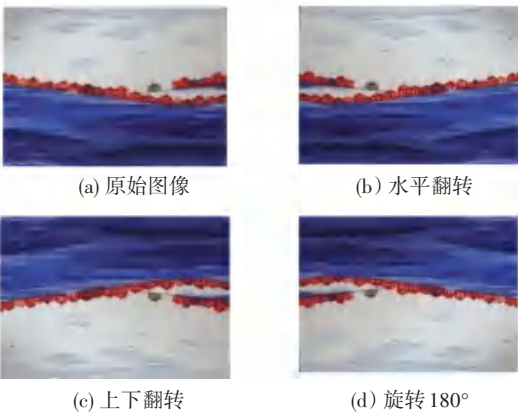


图 7 数据增强示例图

Fig. 7 The schematic figures of data enhancement

4.3 实验评价指标

实验评价指标为图像情感分类准确率,采用 K 折交叉验证方法,因此准确率最终为经 K 折交叉验证后测试集的平均准确率。将标准差 (*Std*) 和 *AUC* (Area Under Curve) 作为实验补充评价指标,通过记录每折测试集的准确率,计算测试集准确率集合的离散程度,得出模型 *Std* 值,而 *AUC* 值为 $[0, 1]$,模型 *AUC* 值越大,证明模型分类的准确率越高。

4.4 实验结果分析

在训练过程中,设置迭代次数为 20,学习率为 0.01, BatchSize 为 64,采用 SGD 优化器,选择 *Crossing - Entropy Loss* 损失函数,动量梯度下降参数设为 0.9,使用 5 折交叉验证方法。为验证本文模型对艺术图像情感识别的效果,选择经典卷积神经

网络模型 ResNet50、ResNet101、ResNet101 与 CBAM 融合算法及李志义提出改进 VGGNet 模型^[15]、Satoshi Iizuka 提出的 ColorNet 中 ClassificationNet^[16] 以及本文提出的特征融合模型进行比较,分析各模型对情感识别准确率变化的情况,5 折交叉验证后测试集的平均准确率见表 1。由表 1 可知,以 ResNet101 与 CBAM 融合特征 X_3 作为主干特征,分别用 $X_1 + X_3$ 、 $X_2 + X_3$ 和 $X_1 + X_2 + X_3$ 3 组特征组合逐一对数据集进行训练,所得的准确率较经典卷积神经网络模型及改进模型相比均有所提升,本文所提出的基于图像底层特征和深层情感语义特征融合的 FeaturesNet 模型准确率可达 93.36%,较 ResNet50 提升 10.31%,较 ResNet101 提升 9.01%,较 ResNet101 与 CBAM 融合模型即 X_3 特征提升 6.98%;与改进模型 VGGNet 相比提升 16.59%;与 ClassificationNet 相比提升 5.31%。从实验结果也可以看出, X_1 与 X_3 的特征组合所得准确率同比 X_2 与 X_3 的特征组合高 0.33%。同时, $X_1 + X_3$ 和 $X_1 + X_2 + X_3$ 模型所对应的标准差较其他模型标准差也较低, *AUC* 的值趋近于 1,说明在测试集上本文模型的准确率表现稳定,波动较小。

表 1 不同模型的准确率对比

Tab. 1 Accuracy comparison with different models

| 模型名称 | 准确率 <i>Acc</i> / % | 标准差 <i>Std</i> | <i>AUC</i> |
|-----------------------------------|--------------------|----------------|------------|
| ResNet50 | 83.05 | 1.36 | 0.853 |
| ResNet101 | 84.35 | 2.35 | 0.815 |
| ResNet101+CBAM (X_3) | 86.38 | 2.94 | 0.829 |
| VGGNet ^[15] | 76.77 | 1.20 | 0.569 |
| ClassificationNet ^[16] | 88.05 | 2.70 | 0.949 |
| $X_1 + X_3$ | 92.86 | 0.52 | 0.975 |
| $X_2 + X_3$ | 92.53 | 1.25 | 0.970 |
| $X_1 + X_2 + X_3$ (FeaturesNet) | 93.36 | 0.55 | 0.979 |

在各模型的 5 折交叉验证中,选取在训练集上预测结果最好的一折,绘制该折训练全过程的准确率及损失值变化情况,各模型的准确率随迭代次数变化的折线对比图如图 8 所示,各模型随迭代次数变化的损失情况如图 9 所示,横坐标代表迭代次数,纵坐标分别代表准确率和损失值。由模型在训练集的表现情况可以得出,特征组合模型 $X_1 + X_3$ 、 $X_2 + X_3$ 和 $X_1 + X_2 + X_3$ 与经典卷积神经网络模型和他人改进模型相比,收敛速度更快,当 *epoch* = 4 时, FeaturesNet 模型的损失值从 4 降至 0.5 以下;当 *epoch* = 6 时, FeaturesNet 模型的损失值已经趋近于

0, 准确率能够达到 100%。

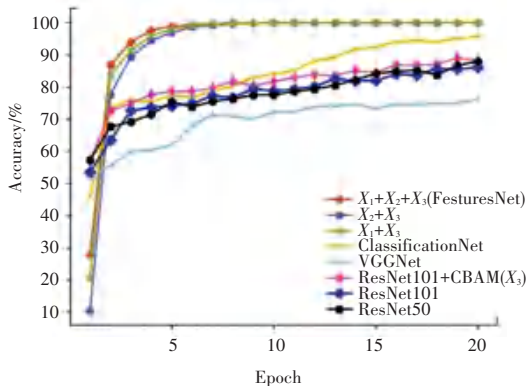


图 8 不同模型的准确率变化折线图

Fig. 8 Line chart of accuracy changes of different models

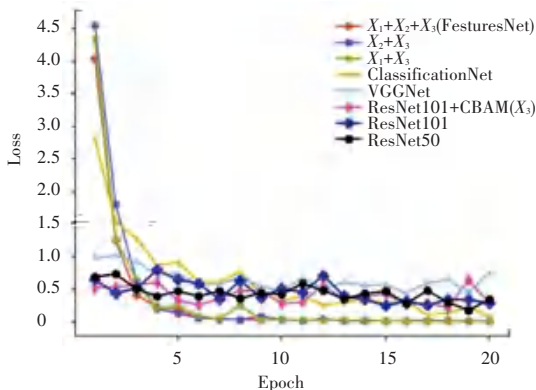


图 9 不同模型的损失值变化折线图

Fig. 9 Line chart of loss value changes of different models

综上,将图像的深度特征与传统特征融合作为艺术图像情感分类模型的输入,能够充分发挥两者对情感分类的重要作用,同时引入基于 CBAM 的注意力机制,使模型的关注点更集中在图像的重点区域,即决定图像情感倾向的区域,从而能够快速有效地提升艺术图像的情感语义分类的准确率。

5 结束语

本文提出一种基于深度特征和注意力机制的卷积神经网络模型 FeaturesNet,将艺术图像的传统底层视觉特征与深层特征相结合,从多角度提取容易激发情感反应的艺术图像特征,可有效解决模型层数的加深带来细节特征丢失的问题,同时引入 CBAM 注意力机制使模型更关注图像中决定情感倾

向的重要区域,从而提升艺术图像情感分类的准确率。实验结果表明,该模型对艺术图像情感分类的准确率可达到 93.36%,证明其对艺术图像情感分类的有效性,为艺术图像的鉴赏与保护工作提供参考。

参考文献

- [1] 汤丽萍,陈芬. 基于情感的图像分类研究进展[J]. 情报理论与实践,2018,41(6):149-153,160.
- [2] 吕微微. 基于情感的图像分类算法研究与实现[D]. 北京:北京交通大学,2012.
- [3] 雷晶晶. 采用模糊聚类和 SVM 的木板套色版画情感语义识别[D]. 昆明:云南大学,2014.
- [4] 李博,郭琛,任慧. 基于加权 K 近邻算法的抽象画图像情感分布预测[J]. 中国传媒大学学报(自然科学版),2018,25(1):36-40.
- [5] KHANNA, PRITEE, YOKOTA, et al. Clustering of Hierarchical Image Database to Reduce Inter- and Intra-semantic Gaps in Visual Space for Finding Specific Image Semantics[J]. Journal of Visual Communication & Image Representation, 2016,38(7):704-720.
- [6] 盛家川,陈雅琦,王君,等. 融合人类认知网络优化的中国画情感识别[J]. 模式识别与人工智能,2020,33(2):141-149.
- [7] 盛家川,陈雅琦,韩亚洪. 深层网络特征聚合重标定的中国画情感分类算法[J]. 计算机辅助设计与图形学学报,2020,32(9):1420-1429.
- [8] 白茹意,郭小英,贾春花. 基于特征融合的小样本抽象画图像情感预测[J]. 计算机应用,2020,40(8):2207-2213.
- [9] 王征,李皓月,许洪山,等. 基于卷积神经网络和 SVM 的中国画情感分类[J]. 南京师大学报(自然科学版),2017,40(3):74-79,86.
- [10] 张浩,徐丹. 基于深度学习的少数民族绘画情感分析方法[J]. 中国科学:信息科学,2019,49(2):204-215.
- [11] 沈哲. 基于图像处理技术的机器人目标跟踪[D]. 赣州:江西理工大学,2010.
- [12] WANG M, WEI Q. Research on the Algorithm of Laplacian Sharpening [C]//Proceedings of the 16th Annual Conference on Information Theory of the Chinese Institute of Electronics. 2009:464-468.
- [13] ITTI L, KOCH C, NIEBUR E. A model of saliency-based visual attention for rapid scene analysis[J]. IEEE Transactions on pattern analysis and machine intelligence, 1998,20(11):1254-1259.
- [14] 董运流. 光场图像的深度估计方法研究[D]. 合肥:合肥工业大学,2019.
- [15] 李志义,许洪凯,段斌. 基于深度学习 CNN 模型的图像情感特征抽取研究[J]. 图书情报工作,2019,63(11):96-107.
- [16] IIZUKA S, SIMO-SERRA E, ISHIKAWA H. Let there be color! Joint end-to-end learning of global and local image priors for automatic image colorization with simultaneous classification[J]. ACM Transactions on Graphics (ToG), 2016,35(4):1-11.