

DOI: 10.11830/ISSN.1000-5013.202410008



# 基于迁移学习和卷积神经网络的 桥梁图像美学评价

叶添照, 赵少杰, 云季彪

(湘潭大学 土木工程学院, 湖南 湘潭 411105)

**摘要:** 为了在桥梁方案设计中实现桥梁美学智能评价,提出一种基于迁移学习和卷积神经网络的桥梁图像美学自动评价方法。首先,通过冻结部分卷积层和修改丢弃率优化 VGG16 网络模型;其次,利用迁移学习将已知数据集 AVA 模型运用到桥梁图像评价上,最终可自动输出对应桥梁美学评分值。结果表明:与人工主观评分相比,文中方法的平均吻合度达到 90.2%,该智能评价方法具有较好的准确性和工程实用性。

**关键词:** 桥梁美学;卷积神经网络;迁移学习;美学评价

中图分类号: TU 026

文献标志码: A

文章编号: 1000-5013(2025)02-0176-07

## Aesthetic Evaluation of Bridge Images Based on Transfer Learning and Convolutional Neural Networks

YE Tianzhao, ZHAO Shaojie, YUN Jibiao

(School of Civil Engineering, Xiangtan University, Xiangtan 411105, China)

**Abstract:** In order to realize intelligent evaluation of bridge schemes in bridge design, an automatic aesthetic evaluation method for bridge images based on transfer learning and convolutional neural networks is proposed. First, the VGG16 network model is optimized by freezing part of convolution layers and modifying the dropout rate. Second, the known data set AVA model is applied to bridge images evaluation by transfer learning, which can automatically output the corresponding aesthetic scores. The results show that compared with the manual subjective evaluation, the average coincidence degree of the proposed method is 90.2%, indicating that the intelligent evaluation method has good accuracy and engineering practicability.

**Keywords:** bridge aesthetics; convolutional neural networks; transfer learning; aesthetic evaluation

随着人工智能技术的日益发展,基于计算机视觉对大量图片进行美学评价变得可能。传统基于人的主观美学评价存在因人而异的缺点,因此,采用计算机视觉技术,摒弃人的主观性,基于图像特征的客观美学质量的评价方法值得深入研究。

目前,基于计算美学方法提取的图像美学特征对图像进行美学分类取得了一些进展。王伟凝等<sup>[1]</sup>提出并行深度卷积神经网络的图像美感分类方法,从多个角度对图片进行特征的自动提取和学习,从而得到更加完整的美感表述。蚁静斌<sup>[2]</sup>通过提取低层视觉特征和高层美学特征,基于机器学习建立图像美感评估分类模型。Sun 等<sup>[3]</sup>提出一套视觉复杂度算子对图像的美学质量进行分类。Zhang 等<sup>[4]</sup>通过多模态递归注意力卷积神经网络(MRACNN)自动提取图像的审美特征,审美质量分类明显优于统一

收稿日期: 2024-10-21

通信作者: 赵少杰(1982-),男,副教授,博士,主要从事桥梁工程安全可靠性的研究。E-mail: shaojiezhao@qq.com。

审美预测任务。文献[5-7]通过精细设计的网络结构,从颜色、布局、多模态综合等角度考虑多种特征因素,以精准地衡量图像的美学价值。目前,利用神经网络对图片进行分类已取得不错的效果,卷积神经网络(CNN)是各种神经网络中最强大的学习结构之一,但鲜少用于桥梁图像的美学评价。梁艳等<sup>[8]</sup>提出桥梁美学应与实用、安全、经济等要素统一考虑。由于图片的复杂性和主观性,局部特征很难全面地表达出图片的要素。李素梅等<sup>[9]</sup>基于卷积神经网络模型对不同舒适度等级的立体图像进行等级分类。王伟凝等<sup>[10]</sup>利用三分法则对图像进行美化。王欣等<sup>[11]</sup>提出多尺度特征提取网络的图像美学可观量化评分方法,利用多尺度特征单元对图片进行深层次挖掘,完成更准确的美学评估。Daichi 等<sup>[12]</sup>构建一个用于美学评价的基于深度学习的卷积神经网络模型,以高精度实现美学评估。Luo 等<sup>[13]</sup>提出一系列构图模板,如矩形分割法、黄金分割法则和当地文化的形态学特征等。Wong 等<sup>[14]</sup>利用视觉注意力机制模型提取了显著区域特征,同时也考虑显著区域和背景之间的联系,得出有区域特征的评价结果更贴合人们的审美评价。Datta 等<sup>[15]</sup>建立了一个较为完善的美学自动评分网站 ACQUINE,实时在线对图片进行打分。Hou 等<sup>[16]</sup>提出基于图像内容的嵌入微调卷积神经网络方法评价图像美学质量,可很好地解决数据规模小的问题。

综上,目前美学评价还很少有关桥梁美学方面的专用模型。现有评价模型由更加广泛的图片组成,所涵盖的特征不能很好地适用于桥梁美学评价。因此,本文结合 VGG16 网络模型,采用迁移学习将学习过的模型运用到桥梁美学评价领域;采用基于特征的迁移,将源域和目标域的数据特征变换到统一特征空间中,再进行桥梁分类识别和美学评价,从而实现从泛化模型到专用模型的转变。

## 1 数据集的构建

深度学习算法依赖于数据样本的全面性和完整性,需要构建适当的桥梁美学图像库,为迁移学习提供合适的训练样本及验证样本,进而通过完成特征学习,实现对桥梁图像的美学评估。

### 1.1 数据集

数据集包括通用数据集和实桥数据集,其中,通用数据集为 Aesthetic Visual Analysis(AVA),用于迁移学习模型的训练;实桥数据集为自定义数据集,用于迁移学习卷积神经网络模型的训练和验证。

1.1.1 AVA 数据集 AVA 数据集包含超过 25 万幅社交图片,每幅图片采集了来自 200 个不同职业、不同年龄的用户评分,评分范围为 1~10 分,以确保评分的客观性和广泛性。

1.1.2 实桥数据集 实桥数据集包括拱桥、梁桥、斜拉桥和悬索桥 4 类。为了使数据集样本充足,对桥梁图片进行裁剪、旋转、平移及灰度处理等操作得到新数据集(图 1),以此产生更多的图片用于后续的特征提取和训练学习。新数据集包含了 12 862 张图片,其中,拱桥有 3 237 张,梁桥有 2 625 张,斜拉桥有 3 500 张,悬索桥有 3 500 张。

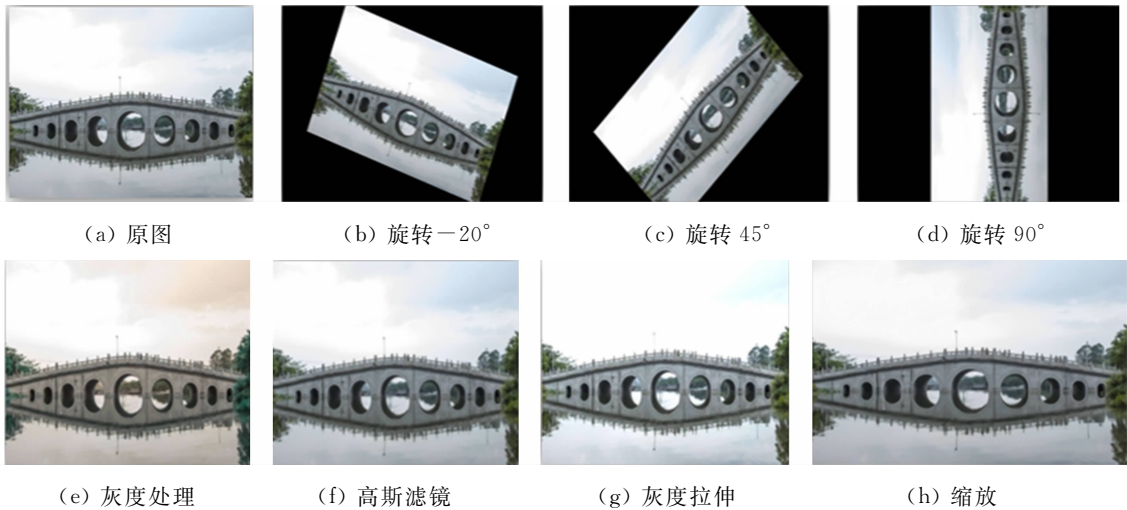


图 1 新数据集

Fig. 1 New dataset

2 基于卷积神经网络的图像美学评估

基于卷积神经网络的图像美学量化方法的步骤如下:1) 数据收集和图像预处理;2) 卷积神经网络模型搭建;3) 模型训练;4) 模型评估与优化;5) 应用模型进行图像美学量化评分。

2.1 卷积神经网络

卷积神经网络作为一种深度学习模型,在图像分类识别上具有显著优势:1) 输入图像和网络的拓扑结构能高度契合;2) 能很好地获取特征图的全局结构信息<sup>[17]</sup>;3) 特征提取和模式分类紧密相连,网络通过多层卷积和池化操作自动学习到有用的特征;4) 权重共享可以减少网络的训练参数,降低模型复杂度,在处理大规模图像数据时具有高效性能。

因此,选取 VGG16 卷积神经网络结构,初始步骤将  $224 \times 224 \times 3$  通道图像输入卷积层。该层利用大小为  $3 \times 3$  的卷积核进行特征提取,卷积步骤的步长设定为 1,卷积层采用“same”填充,以保证特征图的尺寸与输入相同。在卷积过程之后,数据通过 ReLU 激活函数进行非线性处理,随着网络层数的增加,数据分布可能会趋向于饱和极限区域,导致梯度消失。为了解决这一问题,对输入数据进行归一化处理,使其符合均值为 0、方差为 1 的正态分布标准。最后,进入池化层,使用  $2 \times 2$  的池化核进行特征提取,减少特征图的尺寸,同时保留了关键特征信息。经过 4 轮以上的卷积和池化操作后,图像的特征信息得到逐步的提取和精简,桥梁特征信息即可进入全连接层,将特征图的高、宽和通道数转换为一个 4 096 维的向量。这个过程将二维的特征图扁平化为一维,并通过全连接层进行线性 and 非线性变换,最终输出分类结果。

2.2 迁移学习

迁移学习的主要原理是利用源域中已有的大量标注数据,显著减少在新任务上对模型训练所需时间。这种方法在很大程度上增强了模型的稳定性和鲁棒性。其核心目标是将模型在源数据集上积累的先验知识,即模型的权重,转嫁至新的任务或数据集上,以提高学习效率。迁移学习原理,如图 2 所示。迁移学习通常保持模型的卷积层结构不变,已经训练好的权重和参数能够有效地迁移到新的模型结构中。因此,设计适用于新任务冻结部分卷积层的网络模型,为了提高识别准确率,修改丢弃率,其余网络层结构保持不变,并使用新数据集对模型进行训练。实验中将所有层都进行了训练,以保证网络结构的真实性。

2.3 丢弃层

加入丢弃层(dropout),使全连接层节点随机失活,减少模型过拟合现象的出现。具体公式为

$$\mathbf{r}^{(l)} \simeq \text{Bernoulli}(p), \quad \tilde{\mathbf{y}}^{(l)} = \mathbf{r}^{(l)} \mathbf{y}^{(l)}, \quad \mathbf{z}^{(l+1)} = \omega^{(l+1)} \tilde{\mathbf{y}}^{(l)} + b^{(l+1)}, \quad \mathbf{y}^{(l+1)} = f(\mathbf{z}^{(l+1)}). \quad (1)$$

式(1)中:Bernoulli 函数是以概率  $p$  随机生成一个 0、1 的向量; $\tilde{\mathbf{y}}^{(l)}$  表示第  $l$  层随机失活的特征; $\mathbf{y}^{(l)}$ ,  $\mathbf{y}^{(l+1)}$  分别表示第  $l$  层和第  $l+1$  层特征; $\omega^{(l+1)}$ 、 $b^{(l+1)}$  分别表示第  $l+1$  层的权重参数和偏置项; $\mathbf{z}^{(l+1)}$  表示未经激活函数的第  $l+1$  层特征; $f(\mathbf{z}^{(l+1)})$  表示将特征经过激活函数。

2.4 基于迁移学习的 VGG16 网络模型优化

设计优化的基于迁移学习的 VGG16 网络模型,如图 3 所示,总共有 12 层。该网络冻结了 VGG16 模型中第 4 组卷积层的最后 1 层( $28 \times 28 \times 512$ )和第 5 组卷积层( $14 \times 14 \times 512$ ),由 9 个卷积层、3 个全连接层和 2 个丢弃层构成。为了适应特定的分类任务,将模型的修改分类设定为 4。卷积层被分成 4 组,卷积层的卷积核都为  $3 \times 3$ ,步长为 1;池化层的卷积核都为  $2 \times 2$ ,步长为 2。数据第 1 组经过 2 层 64 个卷积核的卷积层捕捉局部特征,尺寸变为  $224 \times 224 \times 64$ ,通道数变为 64,并进行池化处理;第 2 组经过 2 层 128 个卷积核的卷积操作后,尺寸变为  $112 \times 112 \times 128$ ,通道数变为 128,通过池化层;第 3 组经过 3 层 256 个卷积核的卷积操作,尺寸变为  $56 \times 56 \times 256$ ,通道数变为 256,再池化;第 4 组经过 3 层 512

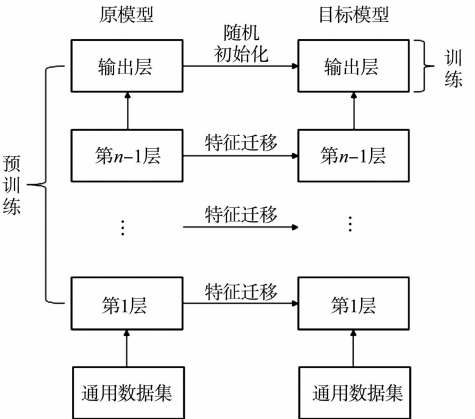


图 2 迁移学习原理  
Fig. 2 Principles of transfer learning

个卷积核的卷积操作,尺寸变为  $28 \times 28 \times 512$ ,通道数变为 512,再进行池化;最后,经过 2 次全连接和丢弃层,再经过一次全连接。

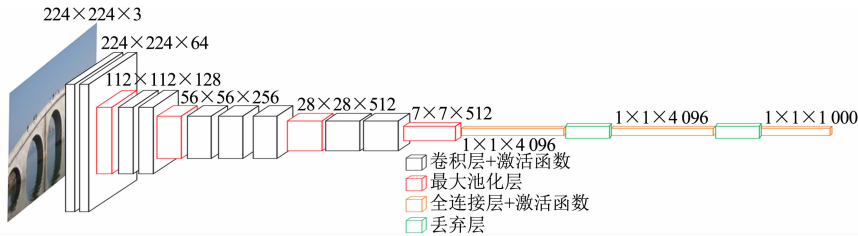


图 3 基于迁移学习的 VGG16 网络模型

Fig. 3 VGG16 network model based on transfer learning

3 实验结果与及分析

测试环境为 Windows 10, GPU 为 NVIDIA GeForce RTX 2060 SUPER, 运用 tensorflow 2.5, python 3.9 进行深度学习。批处理大小设置为 128, 使用 SGD 优化器, 动量参数为 0.9, 初始学习率为 0.01, 每 100 轮训练将学习率下调至原来的 1/10。

3.1 模型验证分析

桥梁分类主要关注准确率( $\eta_a$ )和损失值( $L$ )这两个评估指标, 为确保每种桥型的识别准确率, 修改 dropout 层的丢弃率, 分别设置为 0、0.2、0.4、0.5。不同丢弃率的 VGG16 网络模型实验结果, 如图 4 所示。图 4 中:  $n$  为整个数据集在神经网络中训练过的次数。

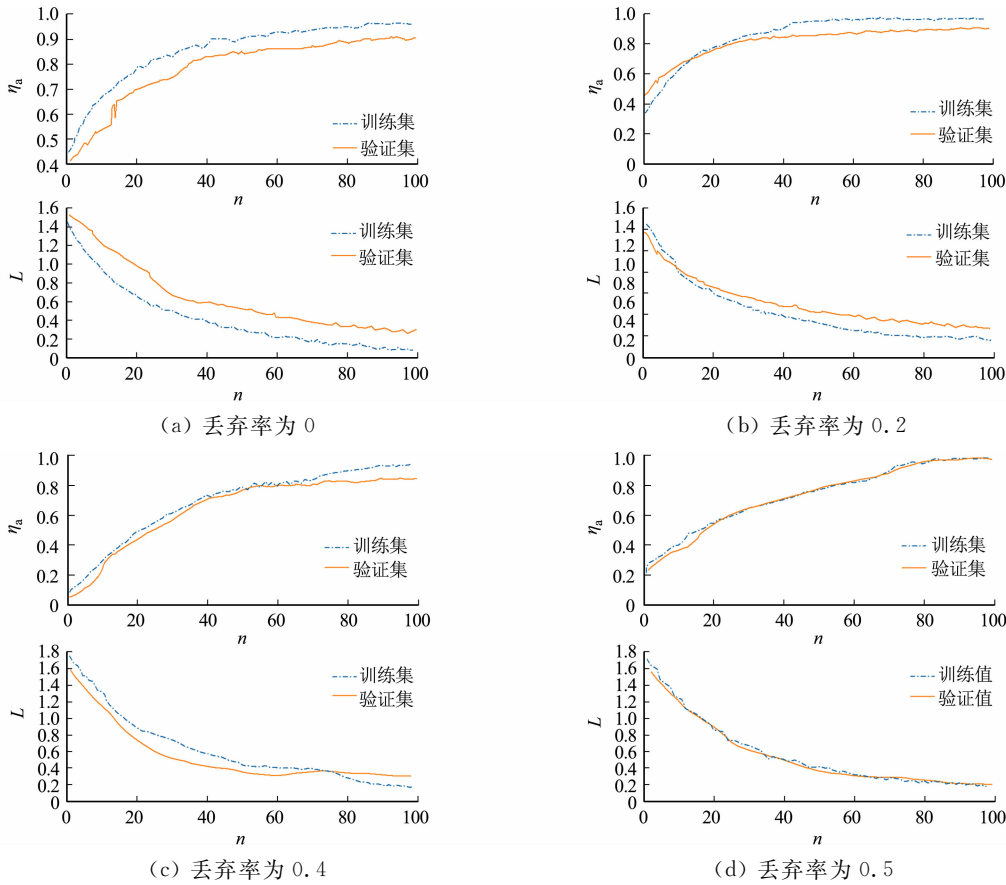


图 4 不同丢弃率的 VGG16 网络模型实验结果

Fig. 4 Experimental results of VGG16 network model with different dropout rates

由图 4(a)可知:模型的准确率和损失在训练集和验证集之间波动幅度较大,训练过程中存在过拟合现象,说明模型在训练数据上表现良好,但在验证数据上表现不佳。由图 4(b)~(d)可知:加入 dropout 层后,模型准确率和损失值的训练集和验证集曲线更加一致,同时减少收敛所需的迭代次数,识

别结果准确率有明显提升;随着丢弃率的增加,准确率和损失值曲线更加接近;当丢弃率为 0.2 时,训练集与验证集的准确率误差为 0.069,训练集与验证集的损失值误差为 0.170;当丢弃率为 0.4 时,训练集与验证集的准确率误差为 0.090,训练集与验证集的损失值误差为 0.150;当丢弃率为 0.5 时,准确率和损失值曲线几乎重合,说明丢弃率为 0.5 的 VGG16 网络模型具有更佳的收敛和拟合效果。

利用已有的预训练模型知识迁移到新的任务上,冻结部分卷积层能很好地保留预训练模型已学习到的有用特征。因此,VGG16 网络模型冻结卷积层第 4 组最后 1 层、第 5 组卷积层和全部冻结卷积层。

冻结部分卷积层,结果如图 5(a)所示。由图 5(a)可知:模型在迭代 40 次左右趋于稳定,表明未出现过拟合和梯度消失等问题,网络结构相对比较稳定。冻结全部卷积层,结果如图 5(b)所示。由图 5(b)可知:模型在训练次数达到 20 次时趋于稳定,收敛于 100%和 83%,表明训练集和验证集相差较大;损失值曲线在训练次数达到 20 次时,开始收敛,最终趋于 0,但验证集的损失值开始逐渐增大到 100%,分析可知该过程出现梯度消失现象,参数无法更新,模型不可用。此外,冻结部分卷积层可以简化模型结构,预期效果较好。

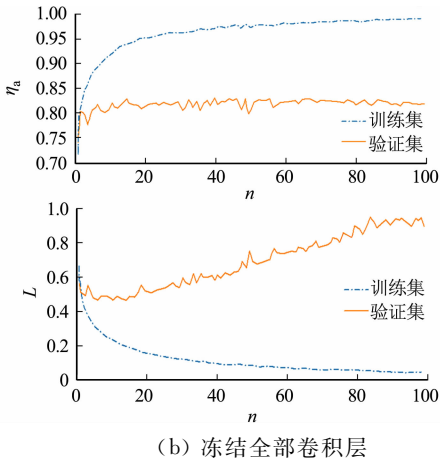
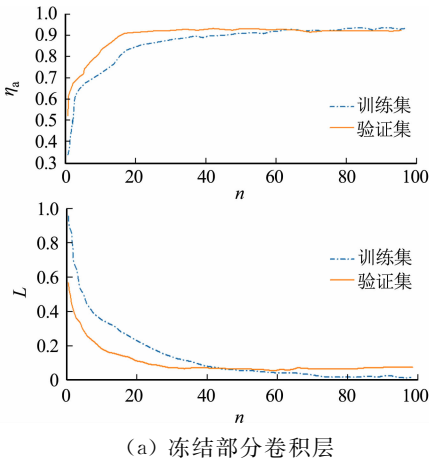


图 5 冻结卷积层实验结果

Fig. 5 Experimental results of frozen convolution layer

按照 8 : 2 划分训练集和验证集,其中,10 290 张为训练集图片,其余为验证集图片。数据集在使用迁移学习的 VGG16 网络模型上的热力图,如图 6 所示。在所建立的数据集中,拱桥识别准确率最高,可以达到 98%;梁桥的识别准确率最低,只有 86%;斜拉桥和悬索桥的识别准确率分别为 94%和 96%。由于拱桥具有独特的拱形结构和实腹式、空腹式等多种构造形式,其拱轴线形状和不同构造形式在图像上呈现出特征多样性。相较于梁桥相对单一的直线梁结构,这些丰富多样的曲线形态和不同构造使得模型能更好地捕捉到区别于梁桥的关键特征,从而使拱桥的识别准确率高于梁桥。由此可见,该数据集有很好的分辨性,且该网络模型也表现出很好的适应性,具有较好的使用性。

文中算法在实桥数据集上的准确率及损失值曲线,如图 7 所示。由图 7 可知:使用迁移学习的模型

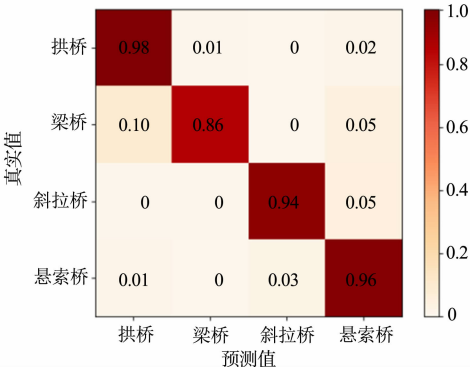


图 6 热力图

Fig. 6 Heatmap

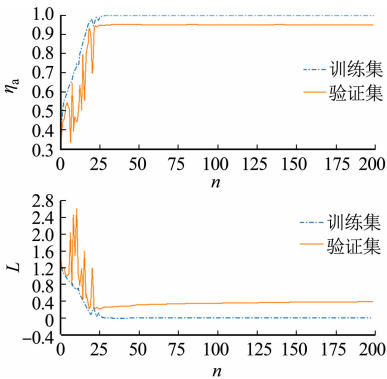


图 7 准确率曲线和损失值曲线

Fig. 7 Accuracy curves and loss value curves



在学习的初步阶段准确率很高,且保持较高上升趋势,但是验证集的准确率和损失值存在较大的振荡,原因是网络模型开始迭代时初始参数随机分布,不同的随机参数会使神经元对输入数据的响应差异较大;由于分类较少,模型在第 30 轮训练后接近收敛,并在后续训练中维持了较高的稳定性,最终准确率达到 94.7%。由此可知,使用迁移学习的网络模型在实验中具有良好表现。

3.2 桥梁图像美学评估

根据以上模型验证分析结果,对 AVA 数据集进行实验。选取 AVA 数据集中评分 1~10 分各分段图像共 4 000 幅,其中,80%作为训练集,20%作为测试集。网络结构模型及参数,如图 3 所示。

基于 AVA 数据集得出实验模型,输入数据集任意桥梁图片。图像美学量化评分结果,如图 8 所示。图 8 中:每幅图片下方的数据为文中方法的结果,括号内数据为人工评分得到的均分;直方图是对应选取 100 人作为调查对象得到的概率密度质量函数图,横坐标为人工评分,纵坐标为对应的概率密度,求均值便得到了人工评分值。由图 8 可知:自动评价结果与人工评分值基本吻合,平均吻合度达到 90.2%,说明使用迁移学习的桥梁图片美学量化方法是准确和可行的。

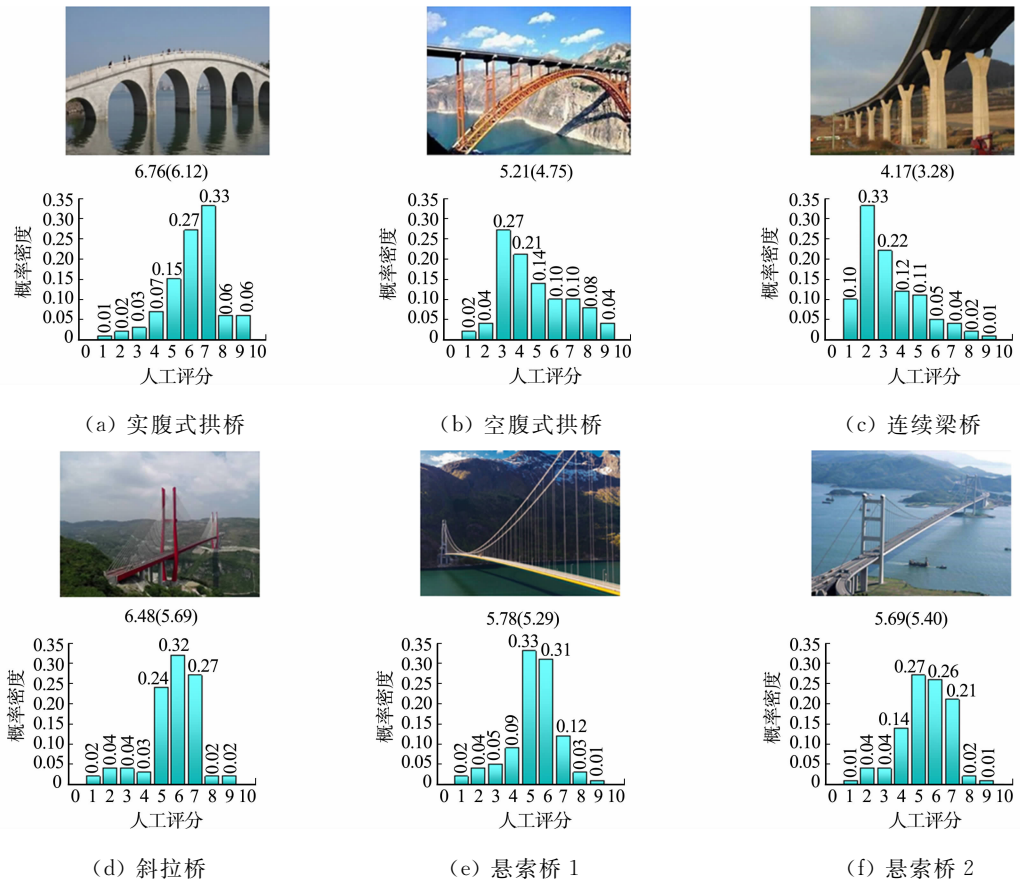


图 8 图像美学量化评分结果

Fig. 8 Quantitative scoring results for image aesthetics

经迁移学习评价桥梁图片的量化方法得到的评分值接近人工评分值,评价精度较好,大部分评分值略高于人工评分值,分析原因主要有以下 2 点:1) 使用迁移学习训练的源数据集包含较多种类,含有的桥梁图片的评分训练较分散,导致结果准确性降低;2) 人工评分因选取的对象专业性不统一,导致分数评价差异较大。因此,后续可进一步优化源数据集,提高人工评分的准确性,从而提高训练网络的评价准确性。

4 结束语

提出一种基于迁移学习和卷积神经网络的桥梁图像美学评估方法,可用于改善桥梁美学评估因数数据量少导致的评估不准确的问题。所用卷积神经网络经过迁移学习优化改善后,采用 AVA 数据集作为源域,冻结部分卷积层和修改丢弃率,输出桥梁图像的客观量化值,从而达到对桥梁图片的评分。实

验结果表明,所提方法能较准确地对桥梁图片进行美学评估。由于数据集的数量限制,所得结果与人工评分值还存在一定的差距,后续工作可进一步提高算法精度,并不断完善数据集。

参考文献:

[1] 王伟凝,王励,赵明权,等. 基于并行深度卷积神经网络的图像美感分类[J]. 自动化学报,2016,42(6):904-914. DOI:10.16383/j. aas. 2016. c150718.

[2] 蚁静绒. 可计算的图像美学分类与评价系统研究[D]. 广州: 华南理工大学,2013.

[3] SUN Litian, YAMASAKI T, AIZAWA K. Photo aesthetic quality estimation using visual complexity features[J]. Multimedia Tools and Applications,2018,77(5):5189-5213. DOI:10.1007/s11042-017-4424-4.

[4] ZHANG Xiaodan,GAO Xinbo,LU Wen,*et al.* Beyond vision: A multimodal recurrent attention convolutional neural network for unified image aesthetic prediction tasks[J]. IEEE Transactions on Multimedia,2020,23:611-623. DOI: 10.1109/TMM. 2020. 2985526.

[5] SHE Dongyu, LAI Yukun, YI Gaoxiong, *et al.* Hierarchical layout-aware graph convolutional network for unified aesthetics assessment[C]//Computer Vision and Pattern Recognition. Nashville:IEEE Press,2021:8471-8480. DOI: 10.1109/CVPR46437. 2021. 00837.

[6] MARTIN-RODRIGUEZ F,GARCIA-MOJON R,FERNANDEZ-BARCIELA M. Detection of AI-created images u- sing pixel-wise feature extraction and convolutional neural networks[J]. Sensors,2023,23(22):9037. DOI:10.3390/ s23229037.

[7] HE Shuai,XIAO Yi,MING Anlong, *et al.* Prompt-guided image color aesthetics assessment: Models, datasets and benchmarks[J]. Information Fusion,2025,114:102706. DOI:10.1016/j. infus. 2024. 102706.

[8] 梁艳,何畏,唐茂林. 桥梁美学 2020 年度研究进展[J]. 土木与环境工程学报(中英文),2021,43(增刊 1):234-241. DOI:10.11835/j. issn. 2096-6717. 2021. 226.

[9] 李素梅,常永莉,段志成. 基于卷积神经网络的立体图像舒适度客观评价[J]. 光学学报,2018,38(6):138-144. DOI: 10.3788/AOS201838. 0610003.

[10] 王伟凝,刘剑聪,徐向民,等. 基于构图规则的图像美学优化[J]. 华南理工大学学报(自然科学版),2015,43(5):51- 58. DOI:10.3969/j. issn. 1000-565X. 2015. 05. 009.

[11] 王欣,穆绍硕,陈华锋. 基于多尺度特征提取网络的图像美学量化评分方法[J]. 浙江大学学报(理学版),2021,48 (1):69-73. DOI:10.3785/j. issn. 1008-9497. 2021. 01. 010.

[12] DAICHI S, HIRONORI T, AKIHIRO K. Study on relationship between composition and prediction of photo aes- thetics using CNN[J]. Cogent Engineering,2022,9(1):2107472. DOI:10.1080/23311916. 2022. 2107472.

[13] LUO Xiaoyu,WU Yue,CHEN Airong, *et al.* Form finding and aesthetic design for pylons of cable-supported bridges[J]. Structural Engineering International,2021,31(6):468-476. DOI:10.1080/10168664. 2020. 1870056.

[14] WONG LAIKUAN,LOW K. Saliency-enhanced image aesthetics class prediction[C]//IEEE International Confer- ence on Image Processing. Cairo:IEEE Press,2009:993-996. DOI:10.1109/ICIP. 2009. 5413825.

[15] DATTA R, JOSHI D, LI Jia, *et al.* Studying aesthetics in photographic images using a computational approach [C]//9th European Conference on Computer Vision. Graz:Springer-Verlag,2006:288-301. DOI:10.1007/1174407 8\_23.

[16] HOU Le, YU Chenping, SAMARAS D. Squared earth mover's distance-based loss for training deep neural net- works[EB/OL]. (2016-11-17)[2024-10-10]. <https://arxiv.org/abs/1611.05916>.

[17] 牛顿,林宁,林振超,等. 多特征融合的焊缝图像多标签分类算法[J]. 华侨大学学报(自然科学版),2024,45(4): 514-523. DOI:10.11830/ISSN. 1000-5013. 202403033.

(责任编辑:黄晓楠      英文审校:陈婧)