

# 绘画艺术图像的计算美学: 研究前沿与展望

鲁越<sup>1,2</sup> 郭超<sup>1,2</sup> 林懿伦<sup>1</sup> 卓凡<sup>3</sup> 王飞跃<sup>1</sup>

**摘要** 绘画艺术是人类艺术创作的重要组成部分, 绘画艺术图像的计算美学是利用机器实现可计算的人类审美过程, 其在大规模绘画的自动化分析和机器对感性的计算建模上具有重要的应用价值和科学意义. 针对其交叉学科的特点, 本文首次从人类审美的感知、认知和评价三个关键过程出发, 将绘画艺术图像的计算美学研究完整地归纳为属性识别、内容理解和美学评价三方面研究内容, 对其中的问题建模、数据获取和前沿方法等关键科学问题进行了归纳总结, 并对绘画计算美学的三方面研究内容进行了对比、思考和展望.

**关键词** 绘画艺术, 计算美学, 审美模型, 属性识别, 内容理解, 美学评价

**引用格式** 鲁越, 郭超, 林懿伦, 卓凡, 王飞跃. 绘画艺术图像的计算美学: 研究前沿与展望. 自动化学报, 2020, 46(11): 2239–2259

**DOI** 10.16383/j.aas.c200358

## Computational Aesthetics of Fine Art Paintings: The State of the Art and Outlook

LU Yue<sup>1,2</sup> GUO Chao<sup>1,2</sup> LIN Yi-Lun<sup>1</sup> ZHUO Fan<sup>3</sup> WANG Fei-Yue<sup>1</sup>

**Abstract** Fine art painting is an essential component of art. The computational aesthetics of fine art painting is a computable human aesthetic process realized by machines, which has significant application value and scientific significance in the automatic analysis of large-scale paintings and computational modeling for aesthetic. Given its interdisciplinary characteristics, for the first time, the computational aesthetics of fine art paintings is completely summarized into three aspects: Attribute recognition, content understanding, and aesthetic judgments according to the key processes of human aesthetics that include perception, cognition, and evaluation. The key scientific issues involved in each aspect are summarized, such as problem modeling, data acquisition, and frontier methods. Also, the three research contents of computational aesthetics of fine art painting are compared, and the future development of this field is discussed.

**Key words** Fine art paintings, computational aesthetics, aesthetic model, attribute recognition, content understanding, aesthetic judgments

**Citation** Lu Yue, Guo Chao, Lin Yi-Lun, Zhuo Fan, Wang Fei-Yue. Computational aesthetics of fine art paintings: The state of the art and outlook. *Acta Automatica Sinica*, 2020, 46(11): 2239–2259

美学是以艺术(特别是绘画艺术)为主要的研究对象, 研究美与丑的审美范畴、人的审美活动的学科<sup>[1]</sup>. 虽然美学与人类感性紧密相关, 但是心理学、神经学等科学方法在美学研究中同样得到了广泛应用. 1876年, 德国美学家费希纳将实验心理学引入到美学研究中, 用定量方法解释视觉刺激, 用客观测量代替主观推理, 创立了实验美学(Experimental aesthetics)<sup>[2]</sup>. 受实验美学的启发, 近年来机器学习也在美学研究中不断探索. 2005年, 计算美学(Computational aesthetics)<sup>[3]</sup>在欧洲图形学会(Euro-

graphics, EG)的第一届图形、视觉和图像的计算美学会议(1st Eurographics Workshop on Computational Aesthetics in Graphics, Visualization and Imaging, CAe 2005)上被提出, 其主要含义是利用机器模仿人类的审美过程, 自主地感知和认知“美”, 并做出美感和情感等美学方面的评价. 从实验美学到计算美学, 美学研究从利用科学方法解释人类审美现象发展为利用机器模仿人类审美过程.

绘画艺术是人类艺术创作的重要组成部分, 是美学在视觉艺术中的主要研究对象. 绘画艺术的计算美学研究具有重要的艺术应用价值和科学价值, 一方面有助于大规模绘画艺术图像的自动化分析, 为艺术欣赏提供辅助信息; 另一方面在机器模仿人类感性行为上做出新探索. 然而, 绘画艺术具有多样的派系和风格以及复杂的表现手法, 其计算美学研究更是涉及机器学习、美学和心理学等交叉学科知识, 充满巨大的挑战.

绘画艺术图像的计算美学目的在于模拟人类的

收稿日期 2020-05-26 录用日期 2020-07-21

Manuscript received May 26, 2020; accepted July 21, 2020

本文责任编辑 魏庆来

Recommended by Associate Editor WEI Qing-Lai

1. 中国科学院自动化研究所复杂系统管理与控制国家重点实验室 北京 100190 2. 中国科学院大学人工智能学院 北京 100049 3. 中央美术学院 北京 100105

1. The State Key Laboratory for Management and Control of Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190 2. School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing 100049 3. Central Academy of Fine Arts, Beijing 100105

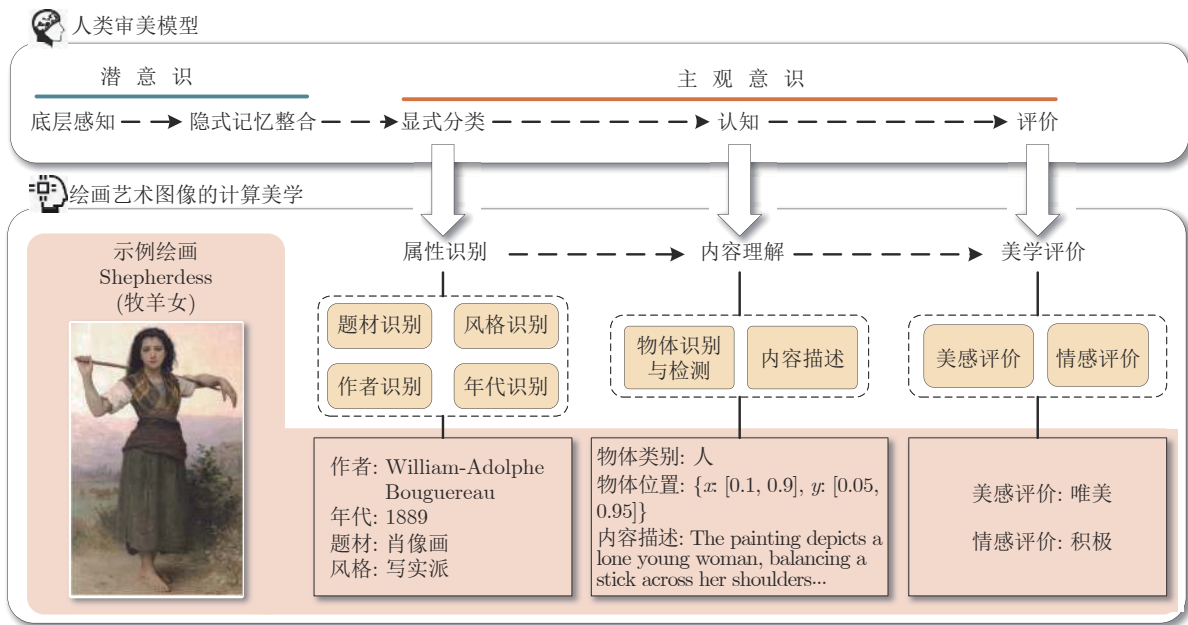


图1 本文对绘画艺术图像计算美学研究的梳理框架及其研究示例<sup>1</sup>

Fig.1 Framework and examples for computational aesthetics of fine art paintings

审美行为. 曾任国际实验美学协会主席的 Leder 将人类的审美行为建模为多层次的信息处理模型, 包含潜意识 and 主观意识两方面过程<sup>[4]</sup>. 潜意识过程包含对颜色、对比度、复杂性等底层信息的感知, 以及对个人经历和记忆的整合, 缺乏显式的信息输出, 难以被量化建模. 而主观意识过程包含显示分类、认知和评价三部分, 具有可被量化的中间结果或审美输出, 可以作为绘画图像计算美学的梳理参考.

为了完整地梳理绘画艺术图像的计算美学研究, 本文参照人类审美模型的主观意识中的显示分类、认知和评价三个过程, 首次将绘画艺术图像的计算美学分为属性识别、内容理解和美学评价三方面研究内容, 图 1 展示了其对应关系及绘画图像计算美学的研究示例.

具体而言, 人类的显式分类是对绘画的初步信息处理, 比如绘画来源和创作背景等信息, 对于计算美学而言这一过程可具体化为属性识别问题, 包括对题材、风格、作者、年代等信息的识别; 人类的认知过程是对绘画的语义和含义的理解, 本文将其和绘画中物体识别与检测及绘画内容描述两个科学问题对应, 构成绘画计算美学的内容理解; 人类的评价过程包含对绘画图像的审美评价和情绪感受, 相应地, 绘画图像的计算美学评价包含美感评价和情感评价两方面问题.

经过关键词检索、引文检索和文献筛选, 本文

<sup>1</sup> 绘画《Shepherdess》图像源于菲尔布鲁克艺术博物馆, 位于美国俄克拉何马州塔尔萨

收集得到近 20 年的绘画艺术图像的计算美学研究共 185 篇文献, 图 2 (a) 展示了文献数量随年份的分布, 可以看出该领域文献数目呈现上升趋势.

为了解该领域内作者合作关系和研究规模, 本文使用社交网络分析 (Social network analysis, SNA)<sup>[5]</sup> 方法, 对所搜集文献的作者合著网络进行分析. 图 2 (b) 展示了利用 Gephi 网络分析软件<sup>[6]</sup> 得到的作者合著网络可视化结果, 其中网络的节点表示作者, 网络的边表示作者间的合作关系, 节点的大小和边的宽度分别表示作者的发文数量及作者间的合作次数. 经统计, 此作者合著网络包含 480 个节点、829 条边、125 个连通性子网, 平均度为 3.45. 从网络的统计数据和可视化效果可以看出, 诸多研究团队进行了相关研究探索, 各个团队子网相对独立, 合作关系较为分散, 这与该领域处于新兴阶段及其多学科特点有关. 图 2 (c) 展示了利用 Gephi 软件对合著网络中规模显著的子网可视化的结果, 作者 Ahmed 所在的团队研究成果相对突出, 包含 7 篇绘画属性识别和美学评价的相关文章. 根据以上作者合著网络的分析, 可以看出该领域研究团队分散, 但初具规模.

近年来, 国内外学者在绘画艺术图像的计算美学相关问题上进行了梳理. 国内方面, 文献 [7] 讨论了自然图像计算美学研究的特征提取、图像构图、图像复杂度等问题; 文献 [8] 总结了绘画图像的实验美学和计算美学的研究方法和评价指标, 然而只讨论了绘画图像的属性分类问题. 国外方面, 文献 [9]

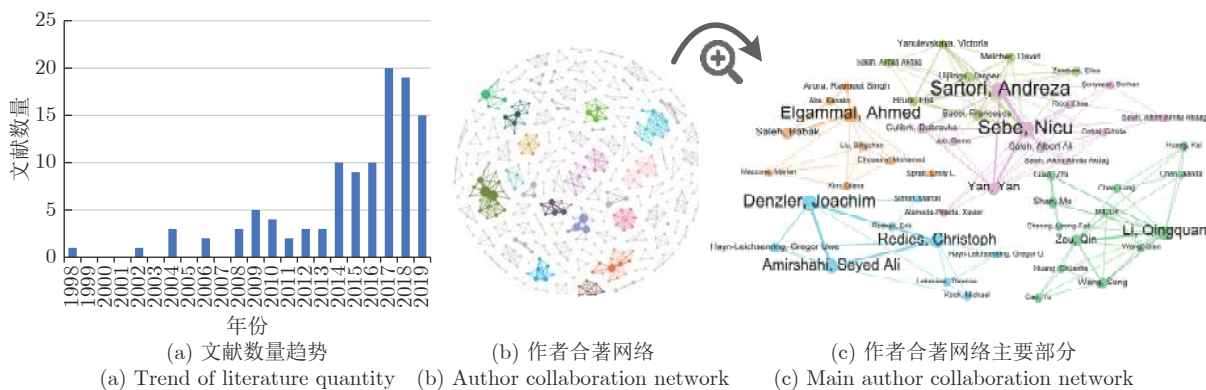


图 2 绘画艺术图像计算美学的文献数量趋势及作者合著网络

Fig.2 Trend of literature quantity and author collaboration network for computational aesthetics of fine art paintings

总结了机器学习在绘画属性识别、赝品鉴别和艺术历史等问题上的研究方法; 文献 [10] 从物理学和数学的角度, 总结了烹饪艺术、绘画艺术、音乐艺术的美感量化评估问题. 以上综述文章只讨论了自然图像或绘画图像的属性识别和美感评估等特定问题, 基于绘画艺术图像的计算美学领域的交叉学科特点, 本文首次从人类的审美过程出发, 完整地梳理了与之对应的属性识别、内容理解和美学评价的计算美学问题, 并探讨它们之间的联系. 文章对其中的问题建模、数据获取、前沿方法等关键科学问题进行了归纳总结, 并对该领域的未来发展做出展望.

第 1 节总结了绘画艺术图像的属性识别, 基于其特征提取方式, 从手工特征和自动特征两方面进行总结; 第 2 节总结了绘画艺术图像的内容理解, 基于其任务类型, 从物体识别与检测、内容描述两方面进行梳理; 第 3 节总结了绘画艺术图像的美学评价, 包含美感评价和情感评价两方面问题; 第 4 节对绘画计算美学三方面研究内容进行了对比、思考和展望; 第 5 节总结全文.

### 1 绘画属性识别

绘画属性识别是绘画计算美学的基本研究内容, 为内容理解和美学评价提供了绘画的基本信息. 绘画属性识别即是由绘画图像判断绘画属性, 包括题材、风格、作者和年代等信息, 图 3(a) 展示了相应的研究示例.

绘画属性识别可以建模为模式分类问题, 针对绘画艺术的图像数据  $x$  与绘画属性  $y$  满足的高维联合分布  $P(x, y)$ , 属性识别即是构建决策函数  $f(x)$  估计其后验概率  $P(y|x)$ , 其核心在于绘画图像特征的提取. 按照特征提取方式的不同, 绘画艺术图像的属性识别可以分为基于手工特征和基于自动特征两种方法.

基于手工特征的方法采用手工特征  $h(\cdot)$  和分类器  $m(\cdot)$  构建决策函数:

$$f(x) = m(h(x)), h(\cdot) \in \mathcal{H} \quad (1)$$

其中  $\mathcal{H}$  是颜色、纹理、结构和高阶语义等特征形成的手工特征集合. 基于手工特征的方法经过手工特征的设计、对比和筛选得到最优的决策函数.

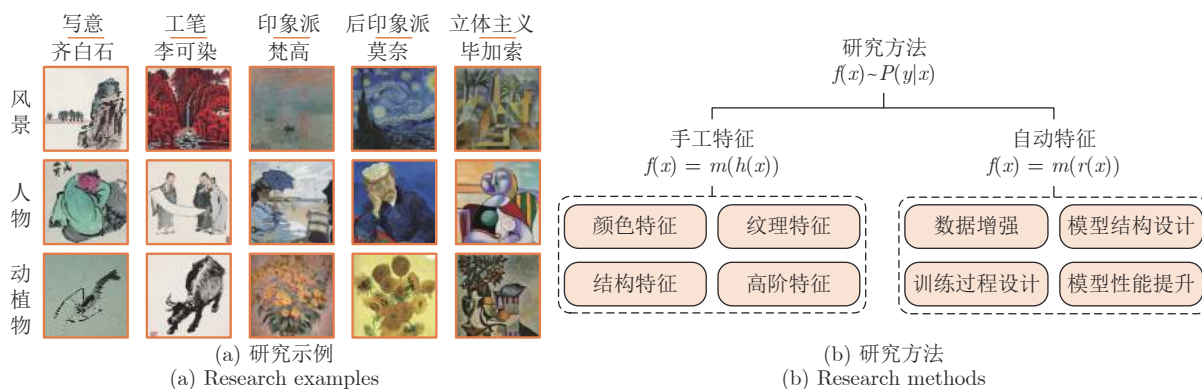


图 3 绘画属性识别任务的研究示例和研究方法

Fig.3 Research examples and methods for attribute recognition of fine art paintings



基于自动特征的方法通过训练自动的特征提取函数  $r(\cdot)$  和分类器  $m(\cdot)$  构建决策函数:

$$f(x) = m(r(x)) \quad (2)$$

基于自动特征的方法一般将特征提取和分类联合训练, 特征提取函数  $r(\cdot)$  由机器学习得到, 典型方法如端到端训练的深度学习方法, 其关键的技术可以归纳为数据增强、模型结构设计、训练过程设计和模型性能提升方法. 图 3(b) 展示了两种绘画艺术图像属性识别方法的关键技术.

### 1.1 基于手工特征的绘画属性识别

基于手工特征的绘画图像属性识别由特征提取和分类器判别两阶段组成, 图 4 展示了其中常用的绘画图像手工特征, 包括颜色特征、纹理特征、结构特征和高阶特征.

颜色是绘画中最直观的表现因素, 艺术家常用色调、冷暖、对比、明暗等颜色的控制方法来表达情感, 传递意义. 此外, 不同的绘画作品和艺术家具有不同的颜色偏好, 这提供了绘画属性识别的区分信息. 文献 [11] 使用颜色矩特征 (Color moment)<sup>[12]</sup> 提取绘画的颜色、色调和色彩范围等信息, 对国画进行山水、花鸟和人物的题材识别以及作者识别. 文献 [13] 利用颜色直方图特征 (Color histogram) 来提取绘画的用色组成信息. 除颜色直方图特征外, 颜色结构描述子特征 (Color structure descriptor, CSD)<sup>[14]</sup> 加入位置信息, 描述颜色的空间分布, 弥补了颜色直方图的缺点; 判别性颜色名称特征 (Discriminative color names, DCN)<sup>[15]</sup> 从信息检索的角度表示了绘画的主要颜色. 文献 [16] 利用多种颜色特征对巴洛克风格、印象派风格和后印象派风格的绘画进行分类, 取得了 80% 以上的正确率.

纹理特征反映了不同的绘画手法, 比如工笔国画相对于写意国画的更精细的纹理. 文献 [17] 利用灰度共生矩阵 (Gray-level co-occurrence matrix, GLCM)<sup>[18]</sup> 来提取绘画艺术图像亮度的局部模式和

排列规则. 与灰度共生矩阵相似, Gram 矩阵<sup>[19]</sup> 在风格迁移研究中用来表征图像风格, 并取得了良好的效果, 是表征绘画手法和笔触风格的良好特征<sup>[20]</sup>. 此外, 频域分解也被用来提取绘画的纹理特征. 小波变换 (Wavelets transform) 利用多尺度的滤波器将图像在频域分解得到其丰富的频域特征, 包括低频信息描述的绘画整体形态, 高频信息描述的绘画纹理和手法. 研究者利用 Haar 小波来提取国画作者间不同的笔法特征<sup>[21]</sup>, 利用 Garbor 小波变换来提取国画的皴法<sup>[22]</sup> 和笔刷特征<sup>[23]</sup>, 利用离散余弦变换 (Discrete cosine transform, DCT) 和边缘特征来鉴别工笔画和写意画<sup>[24]</sup>.

结构特征可以捕捉画面中物体的边缘结构, 进而提供绘画内容上的语义信息. 为了区分立体主义和非立体主义的绘画手法, 文献 [25] 使用 SIFT 特征 (Scale-invariant feature transform)<sup>[26]</sup> 进行绘画属性识别. 原始的 SIFT 特征只能在单通道的灰度图像上计算, 缺少对绘画颜色信息的利用. 经过多通道像素计算的改进和颜色信息的加入, OSIFT 特征 (Opponent SIFT)<sup>[27]</sup> 和 CSIFT 特征 (Color SIFT)<sup>[28]</sup> 在绘画属性识别问题上取得了更好的表现<sup>[29]</sup>. 此外, SURF 特征 (Speeded up robust features)<sup>[30]</sup> 和 ORB 特征 (Oriented FAST and rotated BRIEF)<sup>[31]</sup> 也可以作为绘画的描述特征<sup>[32]</sup>. 局部二值特征 (Local binary patterns, LBP)<sup>[33]</sup> 最初用于人脸识别领域, 可以为肖像画和非肖像画提供良好的区分特征. 另外, 局部亮度顺序特征 (Local intensity order pattern, LIOP)<sup>[34]</sup> 表征局部像素的亮度顺序关系, 向量梯度直方图 (Histogram of oriented gradients, HOG)<sup>[35]</sup> 通过对小块区域中像素梯度方向和强度的直方图统计, 反映对应位置的主要纹理方向, 也是绘画的结构特征提取方法.

高阶特征反映了绘画的物体语义等高阶信息, 通常由低阶特征组合训练的分类器的输出概率构成. GIST 特征<sup>[36]</sup> 反映了图像中的纹理和形状, 有助于区分绘画的内容和场景信息. 此外, Classeme 特

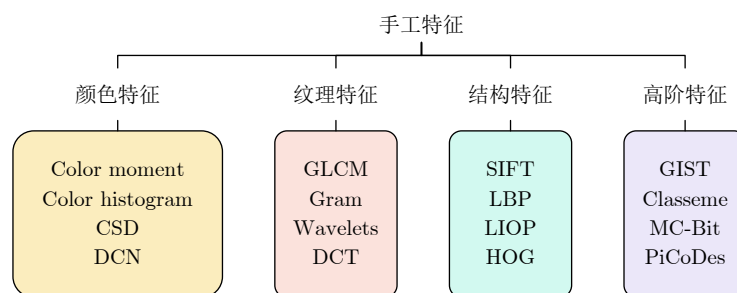


图 4 基于手工特征的绘画属性识别方法的常用特征

Fig.4 Common features for manual features based painting attribute recognition method

征<sup>[37]</sup>、元类别二值特征 (Meta-class binary features, MC-Bit)<sup>[38]</sup> 和 PiCoDes 二值特征<sup>[39]</sup> 均由 HOG、SIFT 和局部自相似度算子 (SSIM)<sup>[40]</sup> 等低阶特征组合训练分类器得到, 对低阶特征进行融合和信息精简, 可以提供与绘画任务相关的更高阶的特征向量. 文献 [41] 利用高阶特征的组合设计了通用的绘画属性识别框架.

对于不同类型的手工特征, 研究者进行了绘画属性识别的对比实验, 我们以风格分类任务为例, 将对比结果汇总在表 1 中. 其中, 文献 [24] 的任务是国画的工笔和写意风格分类, 文献 [42] 和文献 [43] 的任务是西方油画的多风格分类. 虽然各文章使用的数据集存在差异, 但仍然可以看出不同特征间的性能差异和特征性能与任务间的关联性. 边缘特征在文献 [24] 的工笔国画和写意国画分类任务中效果较好, 因为边缘特征更能反映工笔国画和写意国画在手法细腻程度上的差异. SIFT 特征、LBP 特征在文献 [42] 和文献 [43] 的西方油画的风格识别中表现较为突出, 可能因为 SIFT 特征和 LBP 特征捕捉绘画的微观结构信息, 反映了油画绘画的手法和笔触的细微变化.

各种手工特征反映了绘画的风格或手法等不同角度的信息, 一般被筛选、组合来协同使用, 不同特征间也可以相互补足. 比如, 向量梯度直方图特征提供精细的绘画物体识别信息, 但容易受到边缘噪声的影响, 局部二值特征对边缘噪声具有鲁棒性, 弥补了向量梯度直方图的不足. 因此, 不同的特征常利用特征拼接、加权平均、多数投票等方法实现组合协同的使用<sup>[46]</sup>.

某些手工特征在绘画间可能具有不同维度, 比如 SIFT 特征等, 常使用视觉词袋模型 (Bag of visual-

words, BOVW)<sup>[47]</sup> 与其结合, 利用聚类 and 频率分析, 将原始特征转换为频率直方图特征, 解决图像间特征个数不统一的问题. 此外, 为了提升特征表达能力, 常用空间金字塔 (Spatial pyramid)<sup>[48]</sup> 提取的不同尺度特征, 与已有特征结合构成多尺度的特征算子, 比如局部二值特征金字塔 (Pyramidal LBP, PLBP)、向量梯度直方图金字塔 (Pyramid histogram of oriented gradients, PHOG)<sup>[49]</sup> 等.

除了手工特征外, 分类器的选择影响着绘画图像属性识别的性能. 常用的分类器包含朴素贝叶斯分类器、树形分类器 (ID3 决策树、C4.5 决策树<sup>[13]</sup> 和随机森林)、支持向量机、多层感知机和基于  $K$  近邻的聚类分类等方法. 对不同分类器的绘画属性识别性能, 研究者进行了实验对比, 我们以风格识别任务为例, 将结果汇总在表 2 中. 可以看出, 支持向量机和多层感知机在绘画风格识别任务中表现出了较好的性能, 根据文献调研, 二者在基于手工特征的绘画属性识别方法中得到了最为广泛的应用.

整体来说, 在上述手工特征中, 包含颜色特征、纹理特征和结构特征在内的低阶特征提取方法运算简单、鲁棒性强, 在与颜色纹理等低阶特征相关性大的绘画属性识别任务中起到良好的作用, 比如国画工笔和写意手法的识别; 高阶特征融合了多种低阶特征的信息, 提取出语义等高阶的信息, 在与语义相关性强的绘画属性识别任务中起重要作用.

## 1.2 基于自动特征的绘画属性识别

基于自动特征的方法采用自动训练的方式提取特征, 其特征提取与分类一般同时训练, 形成端到端的解决方案. 在绘画属性识别中, 深度卷积网络这种自动的特征方法得到了广泛应用. 本节总结了自动特征方法在数据增强、结构设计、训练过程设计、模

表 1 不同手工特征下的绘画属性识别正确率 (%)

Table 1 Painting attribute recognition accuracy for different manual features (%)

文献	任务	数据集	数据/类别	颜色	HOG	LBP	SIFT	LIOP	小波	GLCM	边缘	GIST
[24]	风格	文章自建	700/2	78.57	83.72	-	80.29	-	81.72	-	86.00	-
[42]	风格	Pandora7k <sup>[42]</sup>	7740/12	-	36.4	52.5	-	36.2	-	-	28.7	33.7
[43]	风格	WikiArt <sup>[41]</sup>	3000/10	36.43	47.97	-	59.20	-	-	35.47	-	39.57

表 2 不同分类器下的绘画属性识别正确率 (%)

Table 2 Painting attribute recognition accuracy for different classifiers (%)

文献	任务	数据集	数据/类别	朴素贝叶斯	树形分类器	支持向量机	多层感知机	$K$ 近邻
[24]	风格	文章自建	700/2	90.29	84.57	95.15	-	82.86
[42]	风格	Pandora7k	7740/12	-	54.00	54.70	-	29.70
[44]	风格	Artchive <sup>[45]</sup>	4119/8	-	63.34	68.51	65.42	-
[23]	风格	文章自建	353/5	48.70	-	57.80	64.00	57.5

型提升上的关键问题和技术,图5展示了本节的讨论框架。

### 1.2.1 数据增强方法

自动特征方法常需要大量的训练数据,而绘画艺术图像的数据一般规模较小,且常具有丰富多样的风格和表现手法。数据增强常用来解决绘画图像训练数据的有限性和数据风格的多样性之间的矛盾,绘画艺术图像的数据增强方法可以分为基本数据增强方法和基于风格迁移的数据增强方法。

基本数据增强方法包含随机裁剪、色彩变换和空间变换等。文献[50]实验了多种基本数据增强方法对少数民族绘画属性识别任务性能的影响,结果表明随机裁剪和颜色、色相、饱和度、对比度等方面的图像微调方法均提升了其任务的正确率。文献[51]考虑到采集绘画图像的相机可能存在不同程度的镜头畸变,将镜头畸变作为一种数据预处理和增强的方法,提高了样本的多样性。基本数据增强方法在应用时需避免破坏与任务相关的图像信息,比如在绘画风格与颜色密切相关时,过大的颜色增强则不适用于绘画风格识别任务。

基于风格迁移的绘画数据增强方法通过将大量的真实图像迁移为绘画风格图像,实现绘画数据的扩充。风格迁移一般利用深度网络特征层的 Gram 矩阵<sup>[19]</sup>表示风格信息,利用深度网络特征层数据本身表示内容信息,通过梯度下降方法不断优化目标图像像素值,使其同时接近自然图像的内容表示和绘画图像的风格表示。随着风格迁移<sup>[19]</sup>算法的发展,

基于风格迁移的数据增强方法成为一种新兴的绘画数据增强手段<sup>[52]</sup>。相比于随机裁剪等传统的图像增强方法,基于风格迁移的数据增强方法引入其他领域的图像内容,显著提高了训练数据的多样性<sup>[52-53]</sup>。虽然基于风格迁移的数据增强生成绘画图像的内容多样,但是生成图像的质量受绘画风格及风格迁移算法等因素制约,生成绘画与真实绘画间仍存在或大或小的数据偏移,因此使用风格迁移实现数据增强时,风格迁移本身也是值得优化的内容,比如调整风格迁移强度系数和生成数据使用比例等参数。

### 1.2.2 模型结构设计

绘画艺术图像的属性信息可能来自图像的局部或全局的视野,而不同的网络结构设置,适用于不同的任务和数据集,模型结构的设计包括基本结构筛选和模型结构优化两个部分。

基本结构筛选利用对比实验确定适应特定绘画任务的模型基础结构。从基本的深度网络 AlexNet<sup>[54]</sup>开始,研究者改进出一系列深度学习模型,比如 VGGNet<sup>[55]</sup>、GoogLeNet<sup>[56]</sup>、InceptionV3<sup>[57]</sup>和 DenseNet<sup>[58]</sup>等。研究者对比了相同绘画图像数据集下各种深度网络结构的性能,本文以风格分类为例,将实验结果整理为表3。可以看出,InceptionV3、ResNet 和 DenseNet 相比于其他网络表现出更优的性能,可能由于在大规模的绘画分类上,残差连接、密集式连接等方式有利于有效绘画属性特征的提取。不同的模型适用于不同的绘画数据集和属性识别任务,因此基本结构筛选是进一步优化前确定基

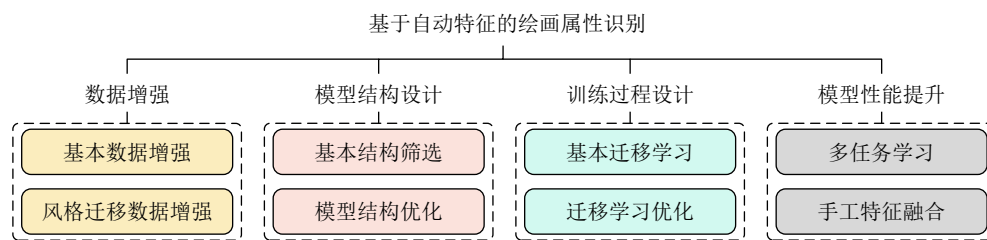


图5 基于自动特征的绘画属性识别方法的关键技术

Fig.5 Common features for automantic features based painting attribute recognition method

表3 不同网络结构下的绘画艺术图像属性识别正确率(%)

Table 3 Painting attribute recognition accuracy for different sturcture of neural networks (%)

文献	任务	数据集	数据/类别	AlexNet	GoogLeNet	InceptionV3	V*-13	V*-16	V*-19	R#-50	R#-101	R#-152	DenseNet
[62]	风格	文章自建	797/17	-	69.90	79.26	-	-	-	76.48	78.43	-	79.36
[46]	风格	WikiArt	30870/6	62.46	64.42	67.16	-	62.69	62.81	66.64	-	-	-
[63]	风格	WikiArt	80000/25	37.80	-	-	-	-	-	49.40	-	-	-
[64]	风格	WikiArt	81449/20	58.20	-	-	60.10	-	-	-	-	63.70	-

V\*: VGG, R#: ResNet



准模型的常用手段。

模型结构优化通过对基准模型的网络结构进行微调, 或融合其他网络结构特点, 以进一步适配绘画艺术图像属性识别任务。文献 [59] 针对国画题材识别任务, 删减了 VGG-16 网络的特征层, 构造出 VGG-15 的网络结构, 将国画题材识别的错误率降低了 8.8%。文献 [60] 将卷积自编码器和卷积神经网络结合, 借助自编码器的信息提取和还原能力, 以提取到更有代表性的绘画特征。

### 1.2.3 训练过程设计

自动特征方法的基本训练方式即利用绘画数据直接训练深度网络, 但是绘画数据的规模可能无法满足深度网络的数据需求。自然图像领域存在大量数据, 训练过程设计的目的在于如何解决自然图像与绘画数据间的数据分布偏差, 通过迁移学习方法利用自然图像数据的知识。下面依次介绍训练过程设计的基本迁移学习方法和迁移学习的优化, 后者包含迁移学习预训练数据集的选择以及迁移学习的权重微调策略。

基本的迁移学习方法利用大规模的自然图像数据对深度网络预训练, 之后在绘画艺术图像数据上微调网络的权重, 其中自然图像一般使用 ImageNet 图像识别任务<sup>[61]</sup>的数据集。文献 [20] 在 WikiArt 数据集上对比随机初始化和迁移学习方式下的绘画艺术图像属性识别性能, 发现经过迁移学习后, 绘画风格、题材和作者识别任务的平均的错误率下降率为 39.8% (错误率下降率 = 错误率变化量/原错误率 × 100%)。从表 4 中的详细数据可以看出, 基本迁移学习方法对绘画题材和作者识别任务的性能提升较大, 对风格识别任务的性能提升相对较小, 这可能是因为在三种任务中, 风格识别任务与自然图像物体识别任务的相关性最低。此外, 在不同的网络

中, ResNet-50、ResNet-98、ResNet-131、DPN-98 和 DPN-131 的迁移学习带来的错误率下降相对较大, 高于平均的错误率下降率。文献 [65] 通过 Visual-BackProp 可视化方法<sup>[66]</sup>对权重微调前后的深度网络观察表明, 微调前的网络激活区域主要集中于表明物体种类的位置, 微调后的网络激活区域移动至与绘画任务更显著相关的位置。

绘画属性识别的迁移学习预训练数据集选择是一种迁移学习的优化方式。根据迁移学习的原理, 两个任务间的相似性越高, 迁移学习的效果越好<sup>[67]</sup>。文献 [68] 研究了不同的自然图像预训练数据集对绘画属性识别性能的影响, 包括 ImageNet 数据集、Places 物体分类数据集<sup>[69]</sup>与 ImageNet 数据集的合并集、LaMem 图片记忆力检测数据集<sup>[70]</sup>、DeepSent 情感识别数据集<sup>[71]</sup>、Flickr 风格识别数据集<sup>[72]</sup>。在 5 种数据集下对绘画属性识别模型预训练, 再基于 WikiArt 数据集针对模型微调, 得到 CaffeNet、HybridNet、LaMemNet、SentimentNet 和 FlickrNet 5 个模型, 表 5 展示了相应的性能指标。根据表 5 的实验结果, HybridNet 和 SentimentNet 网络的性能相对较好, 前者的优势在于大量的训练数据, 而后者的优势可能在于记忆度检测任务与绘画属性识别任务具有更加相关的特性, 使数据集规模较小的情况下仍有较优性能。此外, 文献 [65] 在其他相关绘画数据集上预训练后再利用原绘画数据集微调, 相比于仅利用原绘画数据集, 属性识别的错误率下降了 16.2%。深度网络激活层的可视化表明, 在相关艺术图像数据集上预训练的深度网络, 更能聚焦于有辨别力的图像区域<sup>[65]</sup>。以上分析表明, 迁移学习的预训练数据集规模越大或与绘画属性识别任务的相关性越高, 迁移学习的效果可能越好。

此外, 可以通过迁移学习权重微调策略来优化

表 4 不同初始化方式下的绘画属性识别错误率<sup>[20]</sup> (%)

Table 4 Painting attribute recognition error rate for different initialization methods<sup>[20]</sup> (%)

	风格识别		题材识别		作者识别		平均的错误率下降率
	随机初始化	迁移学习	随机初始化	迁移学习	随机初始化	迁移学习	
AlexNet	69.2	56.7	51.2	35.0	53.7	27.3	33.0
ResNet-14	62.3	51.5	48.7	32.9	44.3	19.6	35.1
ResNet-50	67.2	49.9	51.6	31.0	57.8	18.1	44.7
ResNet-98	69.7	52.1	53.5	31.4	60.9	18.7	45.3
ResNet-131	71.9	53.5	55.2	31.8	65.3	19.9	45.8
DPN-14	54.2	47.8	41.5	27.7	32.8	16.4	31.7
DPN-50	55.4	46.4	43.2	26.3	35.2	16.0	36.6
DPN-98	56.9	44.8	45.0	26.0	36.6	15.6	40.3
DPN-131	60.5	45.0	47.3	25.3	40.4	14.1	45.7
平均的错误率下降率		20.7		38.7		59.9	39.8

表 5 不同预训练数据集下的绘画属性识别的性能  
Table 5 Painting attribute recognition performance for different pre-trained dataset

	CaffeNet	HybridNet	LaMemNet	SentimentNet	FlickrNet
预训练场景	物体分类	物体分类	记忆度检测	乐观度检测	风格分类
预训练数据集	ImageNet	Places + ImageNet	LaMem	DeepSent	FlickrStyle
预训练图片数 (张)	120 万	350 万	6 万	1269	8 万
预训练类别	1000 类	1183 类	[0, 1]*	[0, 1]*	20 类
风格识别正确率 (%)	54.2	56.3	52.6	55.8	50.7
题材识别正确率 (%)	77.2	77.6	75.9	77.4	75.5
作者识别正确率 (%)	76.3	79.1	72.5	78.7	71.4

[0, 1]\*: 0到1的连续等级范围

模型迁移效果. 由于深度网络不同层提取不同等级的特征, 卷积网络的低层主要提取颜色纹理等低阶特征, 高层主要提取与任务相关的高阶语义特征, 不同的微调策略也将产生不同的影响. 文献 [68] 研究了微调过程中冻结低层卷积层的个数对微调后属性识别性能的影响, 实验发现对预训练的 CaffeNet 冻结低层三层卷积层, 将在 WikiArt 数据集上呈现最好的风格分类性能, 相比于微调全部的卷积层, 分类错误率下降了 1.6 %.

### 1.2.4 模型性能提升

在基于自动特征的绘画属性识别方法中, 除数据增强、模型结构设计和训练过程设计外, 还存在其他进一步提升模型性能的方法, 比如多任务学习和手工特征融合方法.

多任务学习通过同时学习多个绘画属性识别任务, 加入更多的关联约束, 进而提取出更显著的绘画特征, 从而提高多个绘画属性识别任务的性能. 绘画的各个属性识别任务之间具有相互关联的特点, 比如微距绘画中常出现花、虫、鸟等景物, 这使绘画风格识别和题材识别的任务间具有相互促进的作用. 深度网络具有连接灵活的神经元结构, 也有利于任务间的合并和多任务模型的搭建. 文献 [73] 在 OmniArt 数据集<sup>[73]</sup>上的实验结果 (表 6) 表明了多任务学习的有效性, 绘画图像的作者识别、类型识别和材质识别三种任务组成的多任务学习, 相比单任务学习错误率下降了 20.13 %. 文献 [74] 不仅利用风格、作者、年代和国别 4 种绘画属性识别任务提高深度网络的特征提取能力, 还通过作品、艺术家、风

格等信息构成的知识网络提取额外的监督信息, 进一步提高网络对绘画上下文信息的提取能力.

手工特征融合将专家知识融入自动特征模型中. 直接的手工特征融合方式是将手工特征与自动特征结合使用, 文献 [75] 在原始深度网络特征的基础上, 计算特征的 Gram 矩阵以提取频域信息, 利用余弦相似度距离构建损失函数, 使绘画属性识别任务的错误率降低了 10.2 %. 文献 [20, 76] 利用灰度共生矩阵获得绘画的笔触信息, 并结合多尺度的层次化特征提取, 将绘画图像和笔触信息共同作为输入训练深度网络. 此外, 手工特征处理方式也得到了借鉴, 文献 [46] 将绘画图像分割为有重叠的区块, 分别经过卷积网络提取特征, 再将各区块的特征拼接, 由多层感知机分类, 相比于全图范围的深度网络判别, 错误率降低了 30.6 %. 这种多区块的特征提取方式缓解了可能出现的误判问题, 增强了模型的鲁棒性. 文献 [63] 在自动特征提取模型中引入 Bagging 机制, 对 WikiArt 数据集进行变换后训练多个深度学习模型, 采用多模型投票的方式确定最终识别结果, 使多模型增强后的绘画风格识别任务的错误率降低了 2.5 %.

### 1.3 数据集及典型方法的性能

本节首先梳理绘画属性识别的主要公开数据集, 包括数据集的规模、特点和来源, 再对比典型的绘画属性识别方法在常用数据集上的性能.

#### 1) 绘画属性识别数据集

按照数据集的规模和标注信息的丰富程度, 公

表 6 单任务与多任务学习的绘画属性识别的性能<sup>[73]</sup> (%)  
Table 6 Painting attribute recognition performance for sigle-task and multi-task learning<sup>[73]</sup> (%)

	作者识别	类型识别	材质识别	平均的错误率下降率
单任务学习错误率	23.3	8.3	2.8	-
多任务学习错误率	21.5	6.3	2.0	-
错误率下降率	7.73	24.10	28.57	20.13



开的绘画属性识别数据集可以分为小规模数据集、大规模数据集和丰富标注数据集三种, 表 7 和表 8 分别展示了它们的基本信息及其具有的标签信息. 在绘画属性识别研究初期, 研究者通过下载绘画图片或扫描书籍中的绘画页以构建绘画属性识别数据集, 形成的数据规模一般较小, 且只具有风格、作者等基本信息的标注, 比如 Painting-91 数据集<sup>[77]</sup> 和 Pandora7k 数据集<sup>[42]</sup>. 随着艺术博物馆的藏品电子化, 出现了开放获取的大规模绘画艺术资料, 研究者对其整理, 形成了大规模的标注数据集, 比如 WikiArt 数据集<sup>[41]</sup>. 在大规模标注数据集中, 有些数据集还标注了基本属性信息之外的其他信息, 给其他的绘画计算美学任务提供了标注数据, 比如 BAM 数据集<sup>[84]</sup> 标注了绘画情绪和内容描述文字, SemArt 数

数据集<sup>[82]</sup> 提供了内容描述文字的标注, iMet2019 数据集<sup>[83]</sup> 和 iMet2020 数据集标注了文化、语义方面的关键词标签.

大规模绘画属性识别数据集的图片和标注主要来自线上的艺术博物馆, 下面对这些艺术博物馆简要介绍. 大都会艺术博物馆 (The Metropolitan Museum of Art) 位于美国纽约, 馆藏近五千年的艺术品, 约 1 500 万件, 其中约 20 万件艺术品得到了电子化采集, 并分别于 2019 年和 2020 年在计算机视觉领域顶级学术会议 CVPR 上举办了艺术品标签识别挑战赛. 荷兰国立博物馆 (The Rijksmuseum) 位于荷兰阿姆斯特丹, 馆藏 500 万件绘画艺术品, 于 2014 年举办艺术品分类挑战赛. 网络艺术博物馆 (The Web Gallery of Art) 于 1996 年创办, 馆藏 3 到

表 7 绘画属性识别任务的公开数据集  
Table 7 Datasets for painting attribute recognition

类型	数据集	文献	年份	来源	总数量	类型
小规模	Painting-91	[77]	2014	-	4.3万	绘画艺术
小规模	Pandora7k	[42]	2016	-	7.7万	绘画艺术
大规模	Pandora18k	[78]	2017	WikiArt	1.8万	绘画艺术
大规模	TICC Printmaking	[79]	2017	荷兰国立博物馆	5.8万	绘画艺术
大规模	WikiArt	[41]	2015	WikiArt	8.1万	绘画艺术
大规模	Rijks2014	[80]	2014	荷兰国内博物馆	11.2万	绘画艺术
大规模	OmniArt	[73]	2017	三个博物馆*	43.2万	绘画艺术及摄影
大规模	Art500k	[81]	2017	三个博物馆*	55.4万	绘画艺术
丰富标注	SemArt	[82]	2018	网络艺术博物馆	2.1万	绘画艺术
丰富标注	iMet2019	[83]	2019	大都会艺术博物馆	15.6万	艺术品
丰富标注	iMet2020	-	2020	大都会艺术博物馆	16.8万	艺术品
丰富标注	BAM	[84]	2017	Behance	2500万	绘画及平面设计等

三个博物馆\*: 包括荷兰国立博物馆、网络艺术博物馆、大都会艺术博物馆

表 8 绘画属性识别数据集的标注信息  
Table 8 Labeling information for painting attribute recognition dataset

类型	数据集	标题	作者	年份	题材	派系	风格	材质	类型	情绪	关键词标签
小规模	Painting91		√				√				
小规模	Pandora7k	√	√	√		√					
大规模	Pandora18k	√	√	√		√					
大规模	TICC Printmaking	√	√	√			√				
大规模	WikiArt	√	√	√			√				
大规模	Rijks2014	√	√	√			√	√			
大规模	OmniArt	√	√					√	√		
大规模	Art500k	√	√	√	√	√	√	√			
丰富标注	SemArt				√						√
丰富标注	iMet2019										√
丰富标注	iMet2020										√
丰富标注	BAM				√			√	√	√	√

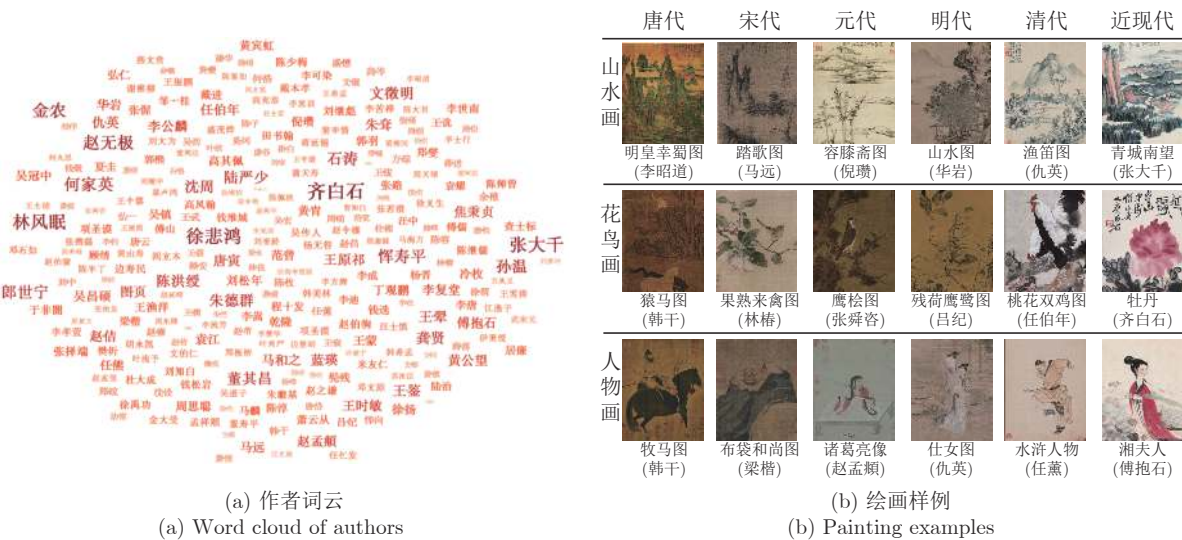


图6 国画属性识别数据库的作者词云和绘画样例

Fig.6 Word cloud of authors and painting examples for Chinese painting attribute recognition database

19 世纪的艺术品图像共约 4.9 万幅, 其中约 2.8 万幅为绘画艺术. 维基艺术 (WikiArt) 是非盈利性质的绘画艺术网站, 拥有 25 万幅绘画图像, 绘画艺术图片由志愿者分享和标注. Behance 平台是一个艺术图片分享网站, 包括艺术家和商业设计师的上传作品, 平台内含有 6500 万的图像数据, 包含雕塑、绘画、摄影、平面设计、涂鸦、插画、广告等多种类型.

在绘画图像属性识别任务的主要公开数据集中, WikiArt 数据集最为常用, WikiArt 数据集是指从维基艺术百科网站<sup>2</sup>下载的绘画数据集, 多个研究对此进行了整理<sup>[85]</sup>. 其中文献<sup>[41]</sup>构建的 WikiArt 数据集比较有代表性, 数据集由 81449 张绘画图片组成, 具有 27 种风格、45 种题材, 包含 1119 个艺术家的作品, 分布从十五个世纪到当代. 数据集中类别间有一定的样本不均衡性, 作品超过 1500 张的题材只有 10 种, 共约 63691 张图像; 作品超过 500 张的艺术家只有 23 个, 共约 18599 张绘画.

由于西方油画图片的开放获取水平较高, 现有的绘画属性数据集中的绘画以西方油画为主. 在中国画方面, 大部分国画属性数据集由网络、书籍或博物馆收集的小规模绘画构成. 天工开源-国画数据库 (OpenSkywork-ChineseClassic Database)<sup>3</sup> 是规模相对较大、比较有代表性的数据库, 包含各年代的国画数据及其属性信息的标注. 图 6 (a) 展示了数据集中包含的画家, 字体大小代表对应画家的画作数量相对多少, 图 6 (b) 以网格形式展示了不同时代的国画样例. 数据库中含有国画图片和年代、作者、

题材等标注信息, 其中的绘画来自 510 个作者, 包括 5771 幅花鸟画、2766 幅山水画、1345 幅人物画、284 幅其他绘画, 总计 10166 幅.

## 2) 典型方法的性能

在本文整理的绘画艺术属性识别的研究中, 大多数文献在 WikiArt 数据集上验证算法性能, 本文统计了典型方法的风格、题材、作者的分类性能 (表 9). 其中, 序号 1~7 为风格分类任务, 8~13 为题材分类任务, 14~18 为作者分类任务. 其中不同方法使用的数据均来自 WikiArt 艺术网站, 因为网站的绘画数量随着时间增多, 不同文献的绘画筛选方法存在差异, 所以算法间的绘画数量和类别数目具有不同.

虽然数据集配置并不完全一致, 但从表 9 中可以看出一些实验结论: 1) 由于 WikiArt 数据集的规模较大, 自动特征方法因其灵活的特征提取能力被广泛使用. 2) 随着数据量的增大, 手工特征以单独特征形式存在较少, 多以融合特征的方式在自动特征方法中加入专家知识. 3) 手工特征与自动特征的融合方法有助于模型性能的提升, 比如序号 6 的方法利用多区域的绘画图像的深度特征进行投票分类, 风格分类正确率较高; 序号 18 的方法利用灰度共生矩阵提取绘画笔触信息, 构建图片和笔触的双通道特征, 实现了较优的绘画作者识别性能.

## 2 绘画内容理解

相比于属性识别, 绘画艺术图像的内容理解是对绘画进一步的语义理解. 具体而言, 绘画艺术图像的内容理解可以分为物体识别与检测、内容描述两方面任务, 前者是对画面中的主要物体进行类识别

2 <https://www.wikiart.org>

3 <https://openskywork.github.io/OpenSkywork-ChineseClassic>

表 9 典型的绘画属性识别方法在 WikiArt 数据集上的性能比较

Table 9 Performance comparison for typical painting attribute recognition methods in WikiArt dataset

序号	任务	年份	文献	方法简介	分类器	数据量	类别数	正确率 (%)
1	风格	2015	[43]	颜色、SIFT、GIST、GLCM	SVM	3000	10	62.37
2	风格	2016	[86]	利用 AlexNet 迁移学习	-	80000	27	54.50
3	风格	2016	[87]	利用 CaffeNet 迁移学习	-	80000	22	55.90
4	风格	2018	[68]	在扩增自然数据集上预训练 ResNet	-	86087	27	56.43
5	风格	2016	[75]	Deep feature、Gram、余弦相似度距离	SVM	82442	25	58.19
6	风格	2019	[46]	由绘画图像块的深度特征经过投票分类	MLP	26400	22	66.71
7	风格	2020	[20]	图片通道和笔触通道形成双通道特征	SVM	30825	25	58.99
8	题材	2016	[86]	利用 AlexNet 迁移学习	-	65000	10	74.14
9	题材	2017	[53]	利用 ResNet 迁移学习	-	79434	26	61.15
10	题材	2018	[68]	在扩增自然数据集上预训练 ResNet	-	96014	10	77.16
11	题材	2015	[43]	颜色、SIFT、GIST、GLCM	SVM	1800	6	84.56
12	题材	2015	[41]	GIST、Claseme、PiCoDes、Deep feature	SVM	63691	10	60.28
13	题材	2020	[20]	图片通道和笔触通道形成双通道特征	SVM	28760	10	76.27
14	作者	2016	[86]	利用 AlexNet 迁移学习	-	20000	23	76.11
15	作者	2017	[88]	利用 ResNet 迁移学习	-	17100	57	77.70
16	作者	2018	[68]	在扩增自然数据集上预训练 ResNet	-	20320	23	81.94
17	作者	2015	[41]	GIST、Claseme、PiCoDes、Deep feature	SVM	18599	23	63.06
18	作者	2020	[20]	图片通道和笔触通道形成双通道特征	SVM	9766	19	88.38

别和包围框定位 (比如人、动物和景物等), 后者是利用自然语言描述画面的主要物体的视觉特点、物体间的关系或正在发生的事件, 图 7 (a) 展示了相应的研究示例。

绘画艺术图像的物体识别与检测问题可以建模成为基于分类的问题, 根据输入图像对物体的类别及包围框坐标进行计算, 目前的方法主要包含弱监督学习和迁移学习方法; 绘画艺术图像的内容描述可以建模成为文字检索问题或文字生成问题, 前者根据输入图像在已有的文字数据库中检索最佳的文字匹配, 后者由输入图像直接生成描述文字。图 7 (b) 展示了相应的关键技术。

根据建模方式的不同, 本节首先分别梳理绘画艺术图像物体识别与检测、内容描述的方法, 最后

介绍常用公开数据集并对比典型方法的性能。

## 2.1 绘画物体识别与检测方法

绘画艺术图像的物体识别与检测的数据集构建需要复杂的人工标注, 因而目前相应的数据集规模较小。为了利用小规模标注数据, 弱监督学习和迁移学习是常用的方法。

弱监督学习利用了无标注数据的信息, 有利于缓解绘画艺术图像物体识别与检测任务标注数据缺少的问题。文献 [89] 使用基于 CycleGAN<sup>[90]</sup> 的风格迁移和伪标签生成两个步骤实现绘画图像的弱监督物体检测。基于 CycleGAN 的风格迁移将带标注的自然图像迁移为绘画风格图像, 同时保留标注信息; 伪标签生成是利用自然图像上预训练的物体检测器生成绘画图像的检测结果, 作为模型微调的伪标签。

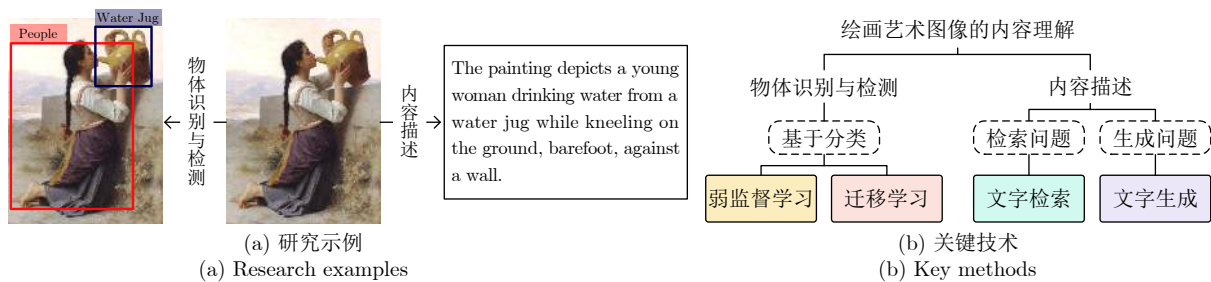


图 7 绘画内容理解的研究示例和关键技术

Fig.7 Research examples and key methods for painting content understanding



迁移学习通过迁移自然图像领域的物体检测知识,帮助绘画艺术图像的物体识别和检测.文献[84]利用 ImageNet 数据集上预训练的 ResNet-50 卷积网络,在绘画数据上微调,实现绘画图像的物体识别.而由于绘画中的物体占据画面的比例可能较小,全图直接物体类别识别可能准确率较低,文献[91]利用自然图像上预训练的目标检测网络 Faster R-CNN<sup>[92]</sup>,直接应用于绘画图像,采用置信度最大的包围框的判别结果作为全图的物体种类识别结果.这种利用物体检测来实现物体识别的方法,缓解了绘画图像中主体比例较小情况下识别困难的问题.在迁移学习的绘画物体识别与检测中,绘画图像与自然图像的数据偏移大小影响着任务的难度.手法抽象的绘画,比如写意手法的中国画或立体主义的西方油画,相比于写实风格的绘画,其物体的颜色、线条、比例与自然图像存在更大差异,带来更大的物体识别与检测难度.文献[93]的研究表明深度学习并不能比较好地泛化到水墨画的物体检测中.对于抽象的绘画作品,相对于鉴赏专家,普通人也只能达到一般的物体检测准确率,比如毕加索作品中的人物检测问题<sup>[94]</sup>.

## 2.2 绘画内容描述方法

绘画艺术图像的内容描述可以建模成为文字检索问题或文字生成问题.

文字检索问题是根据绘画图像从已有的描述文字数据库中检索出相应的最佳匹配.一种可行的方法是将图像和描述编码至同一共享空间,再根据图像与描述的编码距离进行描述检索,在多个研究工作中得到了使用<sup>[82, 95-100]</sup>.文献[96]分别用两个自动编码器分别实现图像和文字描述的编码,利用编码间的余弦相似度距离表示图像和文字的匹配程度.文献[82]还引入了作者、题材、风格和年代信息形成的关键信息网络,促进图像和文字描述的编码.此外,文献[96]还考虑了半监督的场景,无需使用配对的绘画和描述的标注信息,通过最小化最大平均偏差(Maximum mean discrepancy, MMD)来匹配图像和文字间的分布.

将绘画内容描述建模成为文字生成问题是由绘画艺术图像直接生成文字描述.相比于检索问题的建模方式,文字生成问题增加了输出文字的多样性,而不是仅输出数据集中原有的描述,且可以适应新的绘画图像.文献[101]考虑直接由绘画图像生成描述文字的方法,先利用卷积神经网络对绘画图像编码,再经 LSTM 解码,生成对应的文字描述.文献[102]还考虑绘画图像的视觉问答问题(Visual question answering, VQA),算法根据绘画图像及绘画相关的问题输出答案,涉及画面中特定物体个数、颜色、所处地点等信息.

## 2.3 数据集及典型方法的性能

绘画艺术图像内容理解任务的相关数据集需要人工标注,包含物体类别、包围框、文字描述等.不同于绘画艺术图像属性识别,内容理解的标注信息一般不被艺术博物馆的存档信息包含,因此信息的获取更为困难.下面依次对绘画艺术图像的物体识别与检测、内容描述的公开数据集简要介绍.

表 10 展示了绘画艺术图像的物体识别与检测的公开数据集,包括 Paintings 数据集<sup>[103]</sup>、BAM 数据集<sup>[84]</sup>、People-Art 数据集<sup>[104]</sup>、Watercolor2k 数据集<sup>[89]</sup>和神话人物数据集<sup>[105]</sup>,这些数据集均由人工标注获得,前两个数据集只包含图片级别的物体类别标注,其余数据集包含实例级别的物体类别和包围框标注. BAM 数据集是其中规模最大的数据集,包含了约 6.0 万张图片的类别标注信息,但其图片种类混杂,不仅包含绘画图像,还含有大量的平面设计图片.此外, BAM 数据集中每张图片只进行了一次标注,图片标注的准确性参差不齐.神话人物数据集对绘画中的神话人物身份进行了标注.

表 11 展示了绘画艺术图像的内容描述任务的公开数据集,包括 SemArt 数据集<sup>[82]</sup>、EsteArtworks 数据集<sup>[96]</sup>、BibleVSA 数据集<sup>[95]</sup>和 Artpedia 数据集<sup>[97]</sup>,这些数据集的描述语句来自对绘画网站描述和评论的爬取、绘画书籍内容的摘录等,除 SemArt 数据集外,其余数据集的标注信息都经过了人工文字筛选,去除了与绘画无关的描述信息.

表 10 绘画物体识别与检测任务的公开数据集  
Table 10 Datasets for object recognition and detection in paintings

数据集	文献	图片数	类别数	实例数	物体类别标注	物体位置标注
Paintings 数据集	[103]	8 629	10	-	√	
BAM 数据集	[84]	6.0 万	5	-	√	
People-Art	[104]	1 483	1	3 487	√	√
Watercolor2k	[89]	2 000	6	3 315	√	√
神话人物	[105]	2 787	-	-	√	√

表 11 绘画内容描述任务的公开数据集  
Table 11 Datasets for content description of paintings

数据集	文献	图片数目	句子数目	人工核对
SemArt	[82]	21 384	21 384	
EsteArtworks	[96]	553	1 278	√
BibleVSA	[95]	320	2 282	√
Artpedia	[97]	2 930	9 173	√

目前, 绘画艺术图像内容理解的研究刚刚起步, 还没有出现大规模、高质量的标注数据集, 因此, 算法的性能没有统一的对比平台, 表 12 展示了典型的绘画艺术图像内容理解方法及其性能. 在绘画的物体识别与检测中, 常用的评价指标包含准确率和包围框的召回率等, 比如各类别平均准确率的均值 (Mean average precision, mAP). 在绘画的内容描述中, 文字检索方法的常用评价指标是召回率, 比如前 10 个检索结果对应的召回率 (Rcall@10); 描述生成方法的评价指标通过生成文字和标注文字间的相似性计算得到, 常用的评价指标包括 BLEU 评分<sup>[106]</sup>和 CIDEr 评分<sup>[107]</sup>等. 表 12 中的典型方法分别在绘画艺术图像的物体识别、物体检测、描述检索、描述生成、视觉问答 5 种任务上进行了尝试, 为未来的研究提供了参考基准.

### 3 绘画美学评价

相比于属性识别和内容理解, 绘画艺术图像的美学评价是机器对绘画最深度理解, 最接近人类的审美输出. 根据 Leder 的人类审美模型, 人类审美输出包含美感评价和情感评价两方面<sup>[4]</sup>. 同理, 绘画艺术图像的计算美学评价可以分为美感评价和情感评价两方面研究内容.

为了实现机器对人类美学评价的模仿, 首先是对人类的美感和情感进行量化表示, 获得机器可以分析的定量数据. 美感和情感的表示方法可以分为离散词汇表示法和连续变量表示法两种<sup>[108]</sup>. 在离散词汇表示法中, 可以通过气势美、清幽美等词汇表示绘画美感<sup>[109]</sup>, 通过快乐、悲伤、愤怒等词汇表示情感. 在连续变量表示法中, 可以通过从丑到美的连续

等级表示美感, 利用愉悦程度、唤醒程度和支配感等方面连续程度表示情感<sup>[108]</sup>. 人类情感建模的详细综述参见文献 [108].

如果以定量形式表示审美输出, 对于人类而言, 审美的输出  $y_i^{human}$  不仅与绘画艺术图像  $x$  有关, 还与观察者所处环境、初始情绪、审美过程等个人和环境因素有很大关联<sup>[4]</sup>, 将这些因素表示为  $\alpha_i$ , 则观察者的审美输出可以表示为:

$$y_i^{human} = f(x, \alpha_i) \quad (3)$$

由于  $\alpha_i$  不易定量获取, 绘画艺术图像的美学评价研究一般使  $\alpha_i$  变化尽量小, 来利用机器模拟多个观察者在一般环境中的审美共识:

$$y^{machine} = E_{\alpha_i \in \Omega} (f(x, \alpha_i)) \quad (4)$$

其中  $\Omega$  是个人和环境对审美输出的影响因素. 因此, 绘画图像的美学评价的标注信息可以视为多个观察者的平均审美结果, 经过离散词汇表示法或连续变量表示法后划分为离散等级, 绘画图像的美学评价可以建模为模式分类问题.

本节将从美感与情感的评价方法、数据集和典型方法的性能两方面总结绘画美学评价的研究.

#### 3.1 绘画美感与情感的评价方法

基于对已有研究的归纳总结, 目前的美感与情感的评价方法主要分为基于认知心理学的评价方法和基于迁移学习的评价方法.

##### 1) 基于认知心理学的评价方法

认知心理学研究人类的注意、知觉等人类的高级心理过程, 人类的美感和情感体验与人类的认知心理过程紧密相关, 参考人类认知心理学领域的研

表 12 典型的绘画内容理解方法及其性能  
Table 12 Typical painting content understanding methods and performances

文献	任务	方法	数据集	性能
[89]	物体检测	通过风格迁移和伪标签生成进行弱监督学习	Watercolor2k	0.543 (mAP)
[84]	物体识别	利用ResNet-50网络进行迁移学习	BAM 数据集	0.9512 (ACC)
[96]	描述检索	利用自编码器来对齐图像和文字的隐空间	EsteArtworks	0.427 (Rcall@10)
[101]	描述生成	利用卷积网络提取特征, 循环神经网络生成描述	自建数据集	0.970 (CIDEr)
[102]	视觉问答	利用卷积网络提取特征, 循环神经网络生成结果	Artpedia	0.504 (ACC)

究结论,可以对绘画美感与情感的机器评价方法进行设计<sup>[110]</sup>.文献[110]率先进行了绘画美感分类的尝试,根据 Matsuda 的颜色分布理论<sup>[111]</sup>,将绘画的颜色分布匹配到 Matsuda 的颜色理论的 8 种色相类型和 10 种色调类型上,结合其他构图方面的特征,将 100 张梵高和莫奈的印象派风格绘画分类为高美感和低美感两类.

在文献[110]研究基础上,现有的基于认知心理学的评价方法主要从颜色相关的认知理论入手.文献[112]指出颜色与美感程度有较强的相关性,利用单一的颜色特征,进行绘画 4 个美感等级的分类,实现了 73% 的正确率.文献[113]根据心理学领域中颜色与情绪的关联研究<sup>[114]</sup>和艺术家 Itten 的颜色理论<sup>[115]</sup>,选择了颜色、纹理、构图、内容方面的特征,进行绘画情感分类.文献[116]通过艺术层面的理解和构造,进一步设计了用于美感分类的颜色特征,包括颜色自相似度、加权颜色自相似度和颜色异质性等特征.为了排除绘画内容的影响,研究颜色对绘画情感表达的作用, Sartori 进行了一系列的抽象画情感研究,包含色彩和纹理对情感的影响<sup>[117]</sup>、颜色组合对情感的影响<sup>[118]</sup>、绘画基本信息给情感鉴别带来的提高<sup>[119]</sup>以及抽象画的颜色、布局和线条对情感的影响<sup>[120]</sup>.

## 2) 基于迁移学习的评价方法

基于迁移学习的评价方法主要是指迁移自然图像的相关算法和知识到艺术图像的研究中.

在迁移自然图像的相关算法方面,主要借鉴自然图像的手工特征和自动特征等处理算法对绘画进行美学评价.参考自然图像的手工特征,文献[121]使用 SIFT 特征描述子和 LAB 颜色空间的视觉词袋特征进行情感分类;文献[122]提取颜色、纹理和构图的特征,通过浅层神经网络进行情感分类;文献[123]研究了纹理特征对情感分类的重要性.参考自然图像的自动特征提取方法,文献[124]通过 AlexNet 网络提取特征、PCA 降维和 SVM 分类,实现国画的 6 种情感的分类;文献[50]用 VGG-16 网络,借助裁剪、颜色变换等数据增强方法,进行了少数民族绘画情感的积极和消极情绪二分类;文献[125]对比了 VGG-16 模型和 ResNet-15 模型的绘画情感分类性能表现;文献[109]使用树形分类器对国画的美感分类.

在迁移自然图像的知识方面,主要借鉴自然图像的美学数据知识或人类的审美经验知识.借鉴自然图像的美学数据知识时,基本的迁移方式是利用自然图像数据训练模型后,直接应用于绘画图像.文献[126]首次尝试利用在自然图像 IAPS 数据集<sup>[127]</sup>上训练的情感预测模型,直接应用于绘画图像,可以

识别绘画名作的情感,显示出自然图像美学数据知识的可用性.另外,还可以利用自然图像美学分类数据集对深度网络预训练,之后在绘画图像数据集上微调,比如文献[128]在 Twitter DeepSent 数据集<sup>[71]</sup>和 Flickr Sentiment 数据集<sup>[129]</sup>两个自然图像情感数据集上预训练情感识别模型,在 AADB<sup>[130]</sup>、AVA<sup>[131]</sup>、FLICKR-AES<sup>[132]</sup>等自然图像美感数据集上预训练美感识别模型,分别进行绘画图像情感和美感的模型微调和分类.借鉴人类的审美经验知识时,主要考虑多种因素对美学评价的影响,文献[116]首次尝试考虑个人美感品味进行美学评价.文献[133]从情感数据的多维性出发,对绘画图像情感在多维度因素上的分布进行预测.文献[134]考虑了绘画风格、作者对绘画情感的影响,将风格、作者和情感一并建模成为一个多标签的矩阵补全问题.

## 3.2 数据集及典型方法的性能

绘画图像美学数据的标注常通过美学实验获得,比如召集非艺术专业的被试者,在受控的环境和流程下,结合离散词汇和分等级的美学表示方法,记录被试者对绘画图像的审美结果.

作为美学实验的基础,美感和情感的描述词汇的确定过程也需要严谨的理论依据和实验约束.以国画的美感词汇制定为例,文献[109]收集了 350 个国画美感形容词,经过概括审美感受类别的 Hevner 环<sup>[135]</sup>验证词汇完备性,再通过问卷调查的方式筛选出适合国画美感描述的词汇 40 个,之后经过第二轮问卷调查进行词汇的合适度评价,最终筛选出气势美、清幽美、生机美、雅致美和萧瑟美 5 个美感类别,用于国画美感评价.

表 13 展示了公开的绘画美感和情感数据集.由于人类对美感或情感的量化等级敏感度有限,大部分数据集采用 20 个类别和 10 个等级以内的量化表示方法, JenaAesthetics 美感数据集<sup>[109]</sup>包含了最多的 100 个等级的美感程度.由于情感和美感的建模复杂,需要控制场地、人员、流程等多方面变量,相比于绘画的属性识别或内容理解的任务,现有的情感和美感数据库整体规模较小.在表 13 统计的美感和情感数据集中,规模最大的 WikiArt Emotions 绘画情感数据集<sup>[136]</sup>仅包含 4105 张绘画图像.

除了美感和情感的基本标注外,部分数据集还具有其他美学因素的标注,这些标注给深入的美学评价提供了基础. JenaAesthetics 数据集<sup>[137]</sup>包含了颜色、内容、构图上的喜好信息, WikiArt Emotions 数据集<sup>[136]</sup>包含了绘画标题和手法对情感的影响信息. MART 抽象绘画数据集<sup>[121]</sup>中还包含绘画间情感的相对积极程度的标注,令观察者选择给他



表 13 公开的绘画美感和情感数据集  
Table 13 Database for emotion and aesthetic of paintings

类型	数据库	文献	图片数	类别数	等级数	标注/张*	说明
美感	国画美感数据库	[109]	511	5	9	20	气势美、清幽美和生机美上的 9 个等级
美感	JenAesthetics $\beta$	[112]	281	-	4	-	从丑到美 4 个等级
美感	JenaAesthetics	[137]	1629	-	100	20	美感的 100 个等级
情感	国画情感数据库	[109]	511	3	9	20	愉悦度、唤醒度、优势度上的 9 个等级
情感	ArtPhoto 绘画	[113]	807	8	-	-	气愤、激动、害怕等 8 种情感
情感	Affective 抽象绘画#	[113]	228	8	-	14	气愤、激动、害怕等 8 种情感
情感	MART 抽象绘画	[121]	500	-	7	20	消极到积极的 7 个等级
情感	WikiArt Emotions	[136]	4105	20	-	-	害怕、快乐、爱、悲伤等情感

标注/张\*: 每张图片的标注次数; Affective 抽象绘画#: 我们对文献 [113] 的抽象绘画数据集的命名

呈现的两张图中相对积极的图, 之后利用 TrueSkill 排名系统<sup>[138]</sup> 进行绘画的积极程度排序。

表 14 展示了典型的绘画艺术图像的美学评价方法及其性能, 针对美感评价和情感评价共登记了 4 种方法。美学评价方法一般采用正确率来评价算法的美感或情感的分类性能。目前, 绘画艺术图像的美学评价的研究文献数量还较少, 有的研究未给出其方法的性能指标, 比如文献 [113] 的研究。典型的美感和情感评价方法已在小规模的美学数据上取得了一定的效果, 正确率达到 0.75 以上, 未来需要更多的数据标注及方法尝试。

#### 4 思考与展望

在人类的审美过程中, 显示分类、认知和评价三部分相互联系、层层递进。绘画艺术图像的计算美学旨在实现机器对人类审美过程的模拟, 因此本文基于人类审美模型, 将绘画艺术图像的计算美学研究完整地归纳为属性识别、内容理解和美学评价三方面内容, 讨论其中的问题建模、数据获取和前沿方法。相比于人类审美过程的紧密联系性, 绘画艺术图像计算美学的三方面研究相对割裂, 体现在数据、方法和任务三方面。

在数据上, 属性识别、内容理解和美学评价的标注数据集规模存在差异, 其原因在于它们数据获取过程的复杂度不同。属性识别的标注数据由艺术博物馆的藏品信息转化形成, 只需简单的人工整理

与核对; 内容理解的标注数据包括画面的物体类别、包围框和描述文字, 需要大量的人工标注来获取; 美学评价的信息包括美感和情感的量化标注, 需要结合心理学理论设计审美实验来获取标注信息, 过程较为复杂。因此, 绘画属性识别的数据充足, 具有较为成熟的算法性能比较平台, 内容理解和美学评价的数据集仍缺少较大规模的标注数据。此外, 重复的审美活动易引起人类的审美疲劳, 影响标注数据的准确性, 这也是美学评价数据集规模相对较小的原因之一。

在方法上, 属性识别、内容理解和美学评价涉及到的美学和认知心理学知识越来越深入, 相应方法的探索也存在规模上的差异。三个过程与美学的相关程度逐步加深, 涉及到的认知心理学问题也变得更为复杂。比如, 在美学评价研究中, 美感和情感的量化方法及数据审美实验本身就是一个值得深入研究的问题。在本文调查到的 185 篇文献中, 属性识别、内容理解和美学评价的文献数量分别为 120 篇、30 篇和 35 篇, 内容理解和美学评价的研究规模约为属性识别的四分之一。可用数据的规模、问题的复杂度和已有方法的可迁移程度都影响了研究者在属性识别、内容理解和美学评价上的研究规模和探索深度。

在任务联系上, 绘画艺术图像的计算美学的三方面研究还较为割裂, 与人类的审美过程存在差异。人类审美的各个过程间具有信息的传递和反馈, 而

表 14 典型的绘画美学评价方法及其性能  
Table 14 Typical painting aesthetic judgment methods and performances

文献	任务	方法	数据集	图片数/类别数	性能指标 (ACC)
[112]	美感评价	根据颜色特征进行分类	JenAesthetics $\beta$	281/4	0.75
[113]	情感评价	颜色、亮度、纹理等特征	自建抽象画数据集	228/8	-
[116]	情感评价	基于心理学的颜色等特征	MART	500/7	0.78
[109]	情感评价	颜色对比度等特征	国画情感数据库	511/5	0.86

计算美学的三方面研究均是从绘画图像直接判断属性、内容或美学评价的信息, 缺少对其他过程的信息利用. 将属性识别、内容理解和美学评价间的信息联合利用是一个值得探索的方向.

因此, 绘画艺术图像计算美学的研究可以从以下几个方面加强和探索:

#### 1) 构建更为全面和深入的绘画数据集

构建更为全面和深入的绘画数据集需要考虑体量和质量两方面. 在体量上, 由于中西方绘画艺术历史源远流长, 艺术作品丰富多样, 现有的绘画数据集可能只涵盖绘画艺术的一部分, 随着艺术博物馆藏品电子化工作的不断推进, 绘画数据集有望逐步完整. 在质量上, 当前的绘画标注信息主要来自艺术爱好者或机器算法的标注, 其中可能存在噪声信息, 需要绘画艺术专家进行标注信息的核对.

#### 2) 人在回路的计算美学方法

目前的绘画艺术图像计算美学方法中, 人类的监督信息主要以手工标注的形式存在, 没有形成审美信息的反馈闭环, 其结果可能是人类审美的有偏估计, 人在回路的计算美学方法是一个值得探索的方向. 中科院王飞跃团队提出平行艺术<sup>[139-141]</sup>理论体系, 采用基于平行学习<sup>[142]</sup>的多阶段学习方法, 实现人在回路的机器艺术创作和评价. 人在回路的计算美学方法具有人类和机器的动态交互过程, 可能缓解机器相对于人类的审美偏差, 进一步提高机器的审美水平.

#### 3) 计算美学和实验美学的深度融合

当前的计算美学主要从实验美学获取标注数据, 比如美感和情感的标注信息. 然而, 实验美学中丰富的发现和结论没有得到充分的利用, 其原因可能在于研究领域背景知识的差异, 实验美学领域涉及艺术方面的知识和数据的统计分析, 计算美学领域从机器学习角度出发, 研究重点在于模型的设计. 因此, 计算美学和实验美学两个学科的深度融合和专家的密切交流都是未来的可能发展方向.

## 5 结论

绘画艺术图像的计算美学目的在于利用机器模仿人类的审美过程, 在机器的自然图像处理效果不断发展的背景下, 探索机器的感性信息分析能力, 有利于探索人类审美机理和机器的审美能力.

针对绘画艺术图像的计算美学的交叉学科特点, 本文首次从人类审美的主要过程出发, 完整梳理了审美过程中各环节对应的计算美学问题. 本文将绘画艺术图像的计算美学归纳为属性识别、内容理解和美学评价问题, 总结了其中的任务分类、问题

建模, 并深入讨论了前沿方法的原理、数据依赖、性能特点及它们之间的联系. 本文希望提供绘画图像计算美学领域的研究概貌, 为绘画艺术图像的计算美学的整体性、全面性发展提供参考.

## 致谢

作者感谢天工智能(北京)文化科技有限公司的资助.

## References

- Zangwill N. The concept of the aesthetic. *European Journal of Philosophy*, 1998, 6(1): 78-93
- Fechner G T. *Vorschule der Aesthetik*. Leipzig: Breitkopf, 1876. 1-7
- Hoening F. Defining computational aesthetics. In: Proceedings of the 1st Eurographics Conference on Computational Aesthetics in Graphics, Visualization and Imaging. Girona, Spain: Eurographics Association, 2005. 13-18
- Leder H, Belke B, Oeberst A, Augustin D. A model of aesthetic appreciation and aesthetic judgments. *British Journal of Psychology*, 2004, 95(4): 489-508
- Newman M E J. Scientific collaboration networks. I. Network construction and fundamental results. *Physical Review E*, 2001, 64(1): 016131
- Bastian M, Heymann S, Jacomy M. Gephi: An open source software for exploring and manipulating networks. In: Proceedings of the 3rd International Conference on Weblogs and Social Media. San Jose, USA: AAAI, 2009.
- Wang Wei-Ning, Yi Jing-Jian, He Qian-Hua. Review for computational image aesthetics. *Journal of Image and Graphics*, 2012, 17(8): 893-901  
(王伟凝, 蚁静媛, 贺前华. 可计算图像美学研究进展. 中国图象图形学报, 2012, 17(8): 893-901)
- Bai Ru-Yi, Guo Xiao-Ying, Jia Chun-Hua, Geng Hai-Jun. Overview of research methods of painting aesthetics. *Journal of Image and Graphics*, 2019, 24(11): 1860-1881  
(白茹意, 郭小英, 贾春花, 耿海军. 绘画图像美学研究方法综述. 中国图象图形学报, 2019, 24(11): 1860-1881)
- Fiorucci M, Khoroshiltseva M, Pontil M, Traviglia A, Del Bue A, James S. Machine learning for cultural heritage: A survey. *Pattern Recognition Letters*, 2020, 133: 102-108
- Perc M. Beauty in artistic expressions through the eyes of networks and physics. *Journal of The Royal Society Interface*, 2020, 17(164): 20190686
- Liu C, Jiang H. Classification of traditional Chinese paintings based on supervised learning methods. In: Proceedings of the 2014 IEEE International Conference on Signal Processing, Communications and Computing. Guilin, China: IEEE, 2014. 641-644
- Yu H, Li M J, Zhang H J, Feng J F. Color texture moments for content-based image retrieval. In: Proceedings of the 2002 International Conference on Image Processing. Rochester, USA: IEEE, 2002. 929-932
- Jiang S Q, Huang Q M, Ye Q X, Gao W. An effective method to detect and categorize digitized traditional Chinese paintings. *Pattern Recognition Letters*, 2006, 27(7): 734-746
- Manjunath B S, Ohm J R, Vasudevan V V, Yamada A. Color and texture descriptors. *IEEE Transactions on Circuits and Systems for Video Technology*, 2001, 11(6): 703-715

- 15 Khan R, van de Weijer J, Khan F S, Muselet D, Ducottet C, Barat C. Discriminative color descriptors. In: Proceedings of the 2013 IEEE Conference on Computer Vision and Pattern Recognition. Portland, USA: IEEE, 2013. 2866–2873
- 16 Falomir Z, Museros L, Sanz I, Gonzalez-Abril L. Categorizing paintings in art styles based on qualitative color descriptors, quantitative global features and machine learning (QArt-Learn). *Expert Systems with Applications*, 2018, **97**: 83–94
- 17 Lu G M, Gao Z, Qin D N, Zhao X, Liu M J. Content-based identifying and classifying traditional chinese painting images. In: Proceedings of the 2008 Congress on Image and Signal Processing. Sanya, Hainan, China: IEEE, 2008. 570–2574
- 18 Haralick R M. Statistical and structural approaches to texture. *Proceedings of the IEEE*, 1979, **67**(5): 786–804
- 19 Gatys L A, Ecker A S, Bethge M. Image style transfer using convolutional neural networks. In: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE, 2016. 2414–2423
- 20 Zhong S H, Huang X S, Xiao Z J. Fine-art painting classification via two-channel dual path networks. *International Journal of Machine Learning and Cybernetics*, 2020, **11**(1): 137–152
- 21 Sheng Jia-Chuan. Automatic categorization of traditional chinese paintings based on wavelet transform. *Computer Science*, 2014, **41**(2): 317–319  
(盛家川. 基于小波变换的国画特征提取及分类. 计算机科学, 2014, **41**(2): 317–319)
- 22 Li Yu-Fan, Xing Hong-Yan, Chen Jing-Xuan, Yang Min-Zhi. Classification of landscape painting texturing based on gabor. *Computer Science and Application*, 2014, **4**(3): 59–65  
(黎宇帆, 邢鸿雁, 陈静旋, 杨敏之. 基于 Gabor 理论的水画皴法分类. 计算机科学与应用, 2014, **4**(3): 59–65)
- 23 Zujovic J, Gandy L, Friedman S, Pardo B, Pappas T N. Classifying paintings by artistic genre: An analysis of features & classifiers. In: Proceedings of the 2009 IEEE International Workshop on Multimedia Signal Processing. Rio De Janeiro, Brazil: IEEE, 2009. 1–5
- 24 Jiang W, Wang Z, Jin J S, Han Y H, Sun M J. DCT-CNN-based classification method for the gongbi and xieyi techniques of Chinese ink-wash paintings. *Neurocomputing*, 2019, **330**: 280–286
- 25 Misumi M, Orii H, Sharmin T, Mishima K, Tsuruoka T. Image classification for the painting style with SVM. In: Proceedings of the 4th IIAE International Conference on Industrial Application Engineering. Beppu, Japan: The Institute of Industrial Applications Engineers, 2016. 245–249
- 26 Lowe D G. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 2004, **60**(2): 91–110
- 27 Van De Sande K, Gevers T, Snoek C. Evaluating color descriptors for object and scene recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, **32**(9): 1582–1596
- 28 Abdel-Hakim A E, Farag A A. Csfift: A SIFT descriptor with color invariant characteristics. In: Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. New York, USA: IEEE, 2006. 1978–1983
- 29 Saleh B, Abe K, Arora R S, Elgammal A. Toward automated discovery of artistic influence. arXiv: 1408.3218, 2014
- 30 Bay H, Tuytelaars T, Van Gool L. SURF: Speeded up robust features. In: Proceedings of the 9th European Conference on Computer Vision. Graz, Austria: Springer, 2006. 404–417
- 31 Rublee E, Rabaud V, Konolige K, Bradski G. ORB: An efficient alternative to SIFT or SURF. In: Proceedings of the 2011 International Conference on Computer Vision. Barcelona, Spain: IEEE, 2011. 2564–2571
- 32 Haladova Z, Sikudova E. Combination of global and local features for efficient classification of paintings. In: Proceedings of the 29th Spring Conference on Computer Graphics. Smolenice, Slovakia: ACM, 2013. 13–20
- 33 Ojala T, Pietikainen M, Harwood D. A comparative study of texture measures with classification based on featured distributions. *Pattern Recognition*, 1996, **29**(1): 51–59
- 34 Wang Z H, Fan B, Wu F C. Local intensity order pattern for feature description. In: Proceedings of the 2011 International Conference on Computer Vision. Barcelona, Spain: IEEE, 2011. 603–610
- 35 Dalal N, Triggs B. Histograms of oriented gradients for human detection. In: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Diego, USA: IEEE, 2005. 886–893
- 36 Oliva A, Torralba A. Modeling the shape of the scene: A holistic representation of the spatial envelope. *International Journal of Computer Vision*, 2001, **42**(3): 145–175
- 37 Torresani L, Szummer M, Fitzgibbon A. Efficient object category recognition using classemes. In: Proceedings of the 11th European Conference on Computer Vision. Heraklion, Greece: Springer, 2010. 776–789
- 38 Bergamo A, Torresani L. Meta-class features for large-scale object categorization on a budget. In: Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition. Providence, USA: IEEE, 2012. 3085–3092
- 39 Bergamo A, Torresani L, Fitzgibbon A W. PiCoDes: Learning a compact code for novel-category recognition. In: Proceedings of the 24th International Conference on Neural Information Processing Systems. Granada, Spain: NIPS, 2011. 2088–2096
- 40 Shechtman E, Irani M. Matching local self-similarities across images and videos. In: Proceedings of the 2007 IEEE Conference on Computer Vision and Pattern Recognition. Minneapolis, USA: IEEE, 2007. 1–8
- 41 Saleh B, Elgammal A. A unified framework for painting classification. In: Proceedings of the 2015 IEEE International Conference on Data Mining Workshop. Atlantic City, USA: IEEE, 2015. 1254–1261
- 42 Florea C, Condorovici R, Vertan C, Butnaru R, Florea L, Vranceanu R. Pandora: Description of a painting database for art movement recognition with baselines and perspectives. In: Proceedings of the 24th European Signal Processing Conference. Budapest, Hungary: IEEE, 2016. 918–922
- 43 Agarwal S, Karnick H, Pant N, Patel U. Genre and style based painting classification. In: Proceedings of the 2015 IEEE Winter Conference on Applications of Computer Vision. Waikoloa, USA: IEEE, 2015. 588–594
- 44 Condorovici R G, Florea C, Vertan C. Automatically classifying paintings with perceptual inspired descriptors. *Journal of Visual Communication and Image Representation*, 2015, **26**: 222–230
- 45 Cornelis B, Dooms A, Cornelis J, Leen F, Schelkens P. Digital painting analysis, at the cross section of engineering, mathematics and culture. In: Proceedings of the 19th European Signal Processing Conference. Barcelona, Spain: IEEE, 2011. 1254–1258
- 46 Sandoval C, Pirogova E, Lech M. Two-stage deep learning approach to the classification of fine-art paintings. *IEEE Access*, 2019, **7**: 41770–41781



- 47 Csurka G, Dance C R, Fan L X, Willamowski J, Bray C. Visual categorization with bags of keypoints. In: Proceedings of the 8th European Conference on Computer Vision. Prague, Czech Republic: Springer, 2004. 1–22
- 48 Lazebnik S, Schmid C, Ponce J. Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories. In: Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. New York, USA: IEEE, 2006. 2169–2178
- 49 Bosch A, Zisserman A, Munoz X. Representing shape with a spatial pyramid kernel. In: Proceedings of the 6th ACM International Conference on Image and Video Retrieval. Amsterdam, The Netherlands: ACM, 2007. 401–408
- 50 Zhang Hao, Xu Dan. Ethnic painting analysis based on deep learning. *Scientia Sinica Informationis*, 2019, **49**(2): 204–215  
(张浩, 徐丹. 基于深度学习的少数民族绘画情感分析方法. 中国科学: 信息科学, 2019, **49**(2): 204–215)
- 51 Hong Y Y, Kim J. Art painting identification using convolutional neural network. *International Journal of Applied Engineering Research*, 2017, **12**(4): 532–539
- 52 Smirnov S, Eguizabal A. Deep learning for object detection in fine-art paintings. In: Proceedings of the 2018 Metrology for Archaeology and Cultural Heritage. Cassino, Italy: IEEE, 2018. 45–49
- 53 Badea M, Florea C, Florea L, Vertan C. Efficient domain adaptation for painting theme recognition. In: Proceedings of the 2017 International Symposium on Signals, Circuits and Systems. Iasi, Romania: IEEE, 2017. 1–4
- 54 Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 2017, **60**(6): 84–90
- 55 Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv: 1409.1556, 2014
- 56 Szegedy C, Liu W, Jia Y Q, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A. Going deeper with convolutions. In: Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, USA: IEEE, 2015. 1–9
- 57 Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. Rethinking the inception architecture for computer vision. In: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE, 2016. 2818–2826
- 58 Huang G, Liu Z, van der Maaten L, Weinberger K Q. Densely connected convolutional networks. In: Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA: IEEE, 2017. 2261–2269
- 59 Meng Q Y, Zhang H H, Zhou M Q, Zhao S F, Zhou P B. The classification of traditional Chinese painting based on CNN. In: Proceedings of the 4th International Conference on Cloud Computing and Security. Haikou, China: Springer, 2018. 232–241
- 60 David O E, Netanyahu N S. DeepPainter: Painter classification using deep convolutional autoencoders. In: Proceedings of the 25th International Conference on Artificial Neural Networks. Barcelona, Spain: Springer, 2016. 20–28
- 61 Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S A, Huang Z H, Karpathy A, Khosla A, Bernstein M, Berg A C, Li F F. Imagenet large scale visual recognition challenge. *International Journal of Computer Vision*, 2015, **115**(3): 211–252
- 62 Kelek M O, Calik N, Yildirim T. Painter classification over the novel art painting data set via the latest deep neural networks. *Procedia Computer Science*, 2019, **154**: 369–376
- 63 Lecoutre A, Negrevergne B, Yger F. Recognizing art style automatically in painting with deep learning. In: Proceedings of the 9th Asian Conference on Machine Learning. Seoul, Korea: PMLR, 2017. 327–342
- 64 Elgammal A, Mazzone M, Liu B C, Kim D, Elhoseiny M. The shape of art history in the eyes of the machine. arXiv: 1801.07729, 2018
- 65 Sabatelli M, Kestemont M, Daelemans W, Geurts P. Deep transfer learning for art classification problems. In: Proceedings of the 2018 European Conference on Computer Vision. Munich, Germany: Springer, 2018. 631–646
- 66 Bojarski M, Choromanska A, Choromanski K, Firner B, Jackel L, Muller U, Zieba K. Visualbackprop: Efficient visualization of CNNs. arXiv: 1611.05418, 2017
- 67 Pan S J, Yang Q. A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 2010, **22**(10): 1345–1359
- 68 Cetinic E, Lipic T, Grgic S. Fine-tuning convolutional neural networks for fine art classification. *Expert Systems with Applications*, 2018, **114**: 107–118
- 69 Zhou B L, Lapedriza A, Xiao J X, Torralba A, Oliva A. Learning deep features for scene recognition using places database. In: Proceedings of the 27th International Conference on Neural Information Processing Systems. Montreal, Quebec, Canada: NIPS, 2014. 487–495
- 70 Khosla A, Raju A S, Torralba A, Oliva A. Understanding and predicting image memorability at a large scale. In: Proceedings of the 2015 IEEE International Conference on Computer Vision. Santiago, Chile: IEEE, 2015. 2390–2398
- 71 You Q Z, Luo J B, Jin H L, Yang J C. Robust image sentiment analysis using progressively trained and domain transferred deep networks. In: Proceedings of the 29th AAAI Conference on Artificial Intelligence. Austin, USA: AAAI, 2015. 381–388
- 72 Karayev S, Trentacoste M, Han H, Agarwala A, Darrell T, Hertzmann A, Winne-moeller H. Recognizing image style. In: Proceedings of the 2014 British Machine Vision Conference. Nottingham, UK: BMVA, 2014
- 73 Strezoski G, Worring M. OmniArt: Multi-task deep learning for artistic data analysis. arXiv: 1708.00684, 2017
- 74 Garcia N, Renoust B, Nakashima Y. ContextNet: Representation and exploration for painting classification and retrieval in context. *International Journal of Multimedia Information Retrieval*, 2020, **9**(1): 17–30
- 75 Chu W T, Wu Y L. Deep correlation features for image style classification. In: Proceedings of the 24th ACM International Conference on Multimedia. Amsterdam, The Netherlands: ACM, 2016. 402–406
- 76 Huang X S, Zhong S H, Xiao Z J. Fine-art painting classification via two-channel deep residual network. In: Proceedings of the 18th Pacific Rim Conference on Multimedia. Harbin, China: Springer, 2017. 79–88
- 77 Khan F S, Beigpour S, van de Weijer J, Felsberg M. Painting-91: A large scale database for computational painting categorization. *Machine Vision and Applications*, 2014, **25**(6): 1385–1397
- 78 Florea C, Toca C, Gieseke F. Artistic movement recognition by boosted fusion of color structure and topographic description. In: Proceedings of the 2017 IEEE Winter Conference on Applications of Computer Vision. Santa Rosa, USA: IEEE, 2017. 569–577
- 79 van Noord N, Postma E. Learning scale-variant and scale-invariant features for deep image classification. *Pattern Recognition*, 2017, **61**: 583–592
- 80 Mensink T, van Gemert J. The rijksmuseum challenge: Museum-centered visual recognition. In: Proceedings of the 2014 International Conference on Multimedia Retrieval. Glasgow, United Kingdom: ACM, 2014.
- 81 Mao H, Cheung M, She J. DeepArt: Learning joint representations of visual arts. In: Proceedings of the 25th ACM International Conference

- on Multimedia. Mountain View, USA: ACM, 2017. 1183–1191
- 82 Garcia N, Vogiatzis G. How to read paintings: Semantic art understanding with multi-modal retrieval. In: Proceedings of the 2018 European Conference on Computer Vision. Munich, Germany: Springer, 2018. 676–691
- 83 Zhang C Y, Kaeser-Chen C, Vesom G, Choi J, Kessler M, Belongie S. The iMet collection 2019 challenge dataset. arXiv: 1906.00901, 2019
- 84 Wilber M J, Fang C, Jin H L, Hertzmann A, Collomosse J, Belongie S. BAM! The behance artistic media dataset for recognition beyond photography. In: Proceedings of the 2017 IEEE International Conference on Computer Vision. Venice, Italy: IEEE, 2017. 1211–1220
- 85 Saleh B, Elgammal A. Large-scale classification of fine-art paintings: Learning the right metric on the right feature. arXiv: 1505.00855, 2015
- 86 Tan W R, Chan C S, Aguirre H E, Tanaka K. Ceci n'est pas une pipe: A deep convolutional network for fine-art paintings classification. In: Proceedings of the 2016 IEEE International Conference on Image Processing. Phoenix, USA: IEEE, 2016. 3703–3707
- 87 Hentschel C, Wiradarma T P, Sack H. Fine tuning CNNs with scarce training data - adapting imagenet to art epoch classification. In: Proceedings of the 2016 IEEE International Conference on Image Processing. Phoenix, USA: IEEE, 2016. 3693–3697
- 88 Viswanathan N. Artist Identification with Convolutional Neural Networks, Technical Report, SVL Lab, Stanford University, USA, 2017.
- 89 Inoue N, Furuta R, Yamasaki T, Aizawa K. Cross-Domain weakly-supervised object detection through progressive domain adaptation. In: Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE, 2018. 5001–5009
- 90 Zhu J Y, Park T, Isola P, Efros A A. Unpaired image-to-image translation using cycle-consistent adversarial networks. In: Proceedings of the 2017 IEEE International Conference on Computer Vision. Venice, Italy: IEEE, 2017. 2242–2251
- 91 Crowley E J, Zisserman A. In search of art. In: Proceedings of the 2014 European Conference on Computer Vision. Zurich, Switzerland: Springer, 2014.
- 92 Ren S Q, He K M, Girshick R, Sun J. Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, **39**(6): 1137–1149
- 93 Gu Q Q, King R. Deep learning does not generalize well to recognizing cats and dogs in Chinese paintings. In: Proceedings of the 22nd International Conference on Discovery Science. Split, Croatia: Springer, 2019. 166–175
- 94 Ginosar S, Haas D, Brown T, Malik J. Detecting people in cubist art. In: Proceedings of the 2014 European Conference on Computer Vision. Zurich, Switzerland: Springer, 2014. 101–116
- 95 Baraldi L, Cornia M, Grana C, Cucchiara R. Aligning text and document illustrations: Towards visually explainable digital humanities. In: Proceedings of the 2018 International Conference on Pattern Recognition. Beijing, China: IEEE, 2018. 1097–1102
- 96 Carraggi A, Cornia M, Baraldi L, Cucchiara R. Visual-Semantic alignment across domains using a semi-supervised approach. In: Proceedings of the 2018 European Conference on Computer Vision. Munich, Germany: Springer, 2018. 625–640
- 97 Stefanini M, Cornia M, Baraldi L, Corsini M, Cucchiara R. Artpedia: A new visual-semantic dataset with visual and contextual sentences in the Artistic Domain. In: Proceedings of the 20th International Conference on Image Analysis and Processing. Trento, Italy: Springer, 2019. 729–740
- 98 Cornia M, Stefanini M, Baraldi L, Corsini M, Cucchiara R. Explaining digital humanities by aligning images and textual descriptions. *Pattern Recognition Letters*, 2020, **129**: 166–172
- 99 Garcia N, Renoust B, Nakashima Y. Understanding art through multi-modal retrieval in paintings. arXiv: 1904.10615, 2019
- 100 Garcia N, Renoust B, Nakashima Y. Context-aware embeddings for automatic art analysis. In: Proceedings of the 2019 on International Conference on Multimedia Retrieval. Ottawa, Canada: ACM, 2019. 25–33
- 101 Sheng S R, Moens M F. Generating captions for images of ancient artworks. In: Proceedings of the 27th ACM International Conference on Multimedia. Nice, France: ACM, 2019. 2478–2486
- 102 Bongini P, Becattini F, Bagdanov A D, Del Bimbo A. Visual question answering for cultural heritage. arXiv: 2003.09853, 2020
- 103 Crowley E, Zisserman A. The state of the art: Object retrieval in paintings using discriminative regions. In: Proceedings of the 2014 British Machine Vision Conference. Nottingham, UK: BMVA Press, 2014.
- 104 Westlake N, Cai H P, Hall P. Detecting people in artwork with CNNs. In: Proceedings of the 2016 European Conference on Computer Vision. Amsterdam, The Netherlands: Springer, 2016. 825–841
- 105 Madhu P, Kosti R, Muhrenberg L, Bell P, Maier A, Christlein V. Recognizing characters in art history using deep learning. In: Proceedings of the 1st Workshop on Structuring and Understanding of Multimedia HeritAge Contents. Nice, France: ACM, 2019. 15–22
- 106 Papineni K, Roukos S, Ward T, Zhu W J. BLEU: A method for automatic evaluation of machine translation. In: Proceedings of the 40th Annual Meeting on Association for Computational Linguistics. Philadelphia, USA: ACL, 2002. 311–318
- 107 Vedantam R, Lawrence Zitnick C, Parikh D. Cider: Consensus-based image description evaluation. In: Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, USA: IEEE, 2015. 4566–4575
- 108 Li Xia, Lu Guan-Ming, Yan Jing-Jie, Zhang Zheng-Yan. A survey of dimensional emotion prediction by multimodal cues. *Acta Automatica Sinica*, 2018, **44**(12): 2142–2159  
(李霞, 卢官明, 闫静杰, 张正言. 多模态维度情感预测综述. 自动化学报, 2018, **44**(12): 2142–2159)
- 109 Zhan Ying, Gao Yan, Xie Ling-Yun. Feature analysis and classification for aesthetic of Chinese traditional painting. *Journal of Beijing University of Aeronautics and Astronautics*, 2019, **45**(12): 2514–2522  
(湛颖, 高妍, 谢凌云. 中国国画艺术美感特征分析与分类. 北京航空航天大学学报, 2019, **45**(12): 2514–2522)
- 110 Li C C, Chen T. Aesthetic visual quality assessment of paintings. *IEEE Journal of Selected Topics in Signal Processing*, 2009, **3**(2): 236–252
- 111 Matsuda Y. Color Design. Tokyo: Asakura Shoten, 1995. 13–18
- 112 Amirshahi S A, Hayn-Leichsenring G U, Denzler J, Redies C. Color: A crucial factor for aesthetic quality assessment in a subjective dataset of paintings. arXiv: 1609.05583, 2016
- 113 Machajdik J, Hanbury A. Affective image classification using features inspired by psychology and art theory. In: Proceedings of the 18th ACM International Conference on Multimedia. Firenze, Italy: ACM, 2010. 83–92
- 114 Valdez P, Mehrabian A. Effects of color on emotions. *Journal of Experimental Psychology: General*, 1994, **123**(4): 394–409
- 115 Itten J. The Art of Color: The Subjective Experience and Objective Rationale of Color. New York: Wiley, 1997. 15–26
- 116 Amirshahi S A, Denzler J. Judging aesthetic quality in paintings based

- on artistic inspired color features. In: Proceedings of the 2017 International Conference on Digital Image Computing: Techniques and Applications. Sydney, Australia: IEEE, 2017. 1–8
- 117 Sartori A. Affective analysis of abstract paintings using statistical analysis and art theory. In: Proceedings of the 16th International Conference on Multimodal Interaction. Istanbul, Turkey: ACM, 2014. 384–388
- 118 Sartori A, Culibrk D, Yan Y, Sebe N. Who's afraid of itten: Using the art theory of color combination to analyze emotions in abstract paintings. In: Proceedings of the 23rd ACM International Conference on Multimedia. Brisbane, Australia: ACM, 2015. 311–320
- 119 Sartori A, Yan Y, Ozbal G, Salah A A A, Salah A A, Sebe N. Looking at Mondrian's victory Boogie-Woogie: What do I feel? In: Proceedings of the 24th International Joint Conference on Artificial Intelligence. Buenos Aires, Argentina: AAAI, 2015. 2503–2509
- 120 Sartori A, Culibrk D, Yan Y, Job R, Sebe N. Computational modeling of affective qualities of abstract paintings. *IEEE MultiMedia*, 2016, **23**(3): 44–54
- 121 Yanulevskaia V, Uijlings J, Bruni E, Sartori A, Zamboni E, Bacci F, Melcher D, Sebe N. In the eye of the beholder: Employing statistical analysis and eye tracking for analyzing abstract paintings. In: Proceedings of the 20th ACM International Conference on Multimedia. Nara, Japan: ACM, 2012. 349–358
- 122 Leong J, Chung A. Affective Classification of Abstract Paintings Using Artificial Neural Networks, Technical Report, University of Waterloo, Canada, 2014.
- 123 Sartori A, Snayzar B, Salah A A A, Salah A A, Sebe N. Emotions in abstract art: Does texture matter? In: Proceedings of the 18th International Conference on Image Analysis and Processing. Genoa, Italy: Springer, 2015. 671–682
- 124 Wang Zheng, Li Hao-Yue, Xu Hong-Shan, Sun Mei-Jun. Chinese painting emotion classification based on convolution neural network and SVM. *Journal of Nanjing Normal University (Natural Science Edition)*, 2017, **40**(3): 74–79  
(王征, 李皓月, 许洪山, 孙美君. 基于卷积神经网络和 SVM 的中国画情感分类. 南京师大学报 (自然科学版), 2017, **40**(3): 74–79)
- 125 Tan W, Wang J, Wang Y, Lewis M, Jarrold W, Davis U. CNN Models for Classifying Emotions Evoked by Paintings, Technical Report, SVL Lab, Stanford University, USA, 2018.
- 126 Yanulevskaia V, van Gemert J C, Roth K, Herbold A K, Sebe N, Geusebroek J M. Emotional valence categorization using holistic image features. In: Proceedings of the 15th IEEE International Conference on Image Processing. San Diego, USA: IEEE, 2008. 101–104
- 127 Lang P J. International Affective Picture System (IAPS): Technical Manual and Affective Ratings, Technical Report, NIMH Center for the Study of Emotion and Attention, University of Florida, USA, 1997.
- 128 Cetinic E, Lipic T, Grgic S. A deep learning perspective on beauty, sentiment, and remembrance of art. *IEEE Access*, 2019, **7**: 73694–73710
- 129 Katsurai M, Satoh S. Image sentiment analysis using latent correlations among visual, textual, and sentiment views. In: Proceedings of the 2016 IEEE International Conference on Acoustics, Speech and Signal Processing. Shanghai, China: IEEE, 2016. 2837–2841
- 130 Kong S, Shen X H, Lin Z, Mech R, Fowlkes C. Photo aesthetics ranking network with attributes and content adaptation. In: Proceedings of the 14th European Conference on Computer Vision. Amsterdam, The Netherlands: Springer, 2016. 662–679
- 131 Murray N, Marchesotti L, Perronnin F. AVA: A large-scale database for aesthetic visual analysis. In: Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition. Providence, USA: IEEE, 2012. 2408–2415
- 132 Ren J, Shen X H, Lin Z, Mech R, Foran D J. Personalized image aesthetics. In: Proceedings of the 2017 IEEE International Conference on Computer Vision. Venice, Italy: IEEE, 2017. 638–647
- 133 Li Bo, Guo Chen, Ren Hui. Emotion distribution prediction of abstract painting image based on weighted KNN algorithm. *Journal of Communication University of China (Science and Technology)*, 2018, **25**(1): 36–40  
(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. 中国传媒大学学报 (自然科学版), 2018, **25**(1): 36–40)
- 134 Alameda-Pineda X, Ricci E, Yan Y, Sebe N. Recognizing emotions from abstract paintings using non-linear matrix completion. In: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE, 2016. 5240–5248
- 135 Hevner K. Experimental studies of the elements of expression in music. *The American Journal of Psychology*, 1936, **48**(2): 246–268
- 136 Mohammad S, Kiritchenko S. WikiArt emotions: An annotated dataset of emotions evoked by art. In: Proceedings of the 11th International Conference on Language Resources and Evaluation. Miyazaki, Japan: ELRA, 2018.
- 137 Amirshahi S A, Redies C, Denzler J. How self-similar are artworks at different levels of spatial resolution? In: Proceedings of the 2013 Symposium on Computational Aesthetics. Anaheim, USA: ACM, 2013. 93–100
- 138 Herbrich R, Minka T, Graepel T. TrueSkill™: A bayesian skill rating system. In: Proceedings of the 19th International Conference on Neural Information Processing Systems. Vancouver, Canada: MIT, 2006. 569–576
- 139 Wang Fei-Yue. Parallel Art: From Intelligent Art to Artistic Intelligence, Technical Report, The Alfred North Whitehead College, Beijing, China, 2017.
- 140 Guo Chao, Lu Yue, Lin Yi-Lun, Zhuo Fan, Wang Fei-Yue. Parallel art: Artistic creation under human-machine collaboration. *Chinese Journal of Intelligent Science and Technology*, 2019, **1**(4): 335–341  
(郭超, 鲁越, 林懿伦, 卓凡, 王飞跃. 平行艺术: 人机协作的艺术创作. 智能科学与技术学报, 2019, **1**(4): 335–341)
- 141 Guo C, Bai T X, Lu Y, Lin Y L, Xiong G, Wang X, et al. Skywork-davinci: A novel CPSS-based painting support system. In: Proceedings of the 16th International Conference on Automation Science and Engineering. Hong Kong, China: IEEE, 2020. 673–678
- 142 Li Li, Lin Yi-Lun, Cao Dong-Pu, Zheng Nan-Ning, Wang Fei-Yue. Parallel learning — a new framework for machine learning. *Acta Automatica Sinica*, 2017, **43**(1): 1–8  
(李力, 林懿伦, 曹东璞, 郑南宁, 王飞跃. 平行学习—机器学习的一个新型理论框架. 自动化学报, 2017, **43**(1): 1–8)



鲁越 中国科学院自动化研究所复杂系统管理与控制国家重点实验室博士研究生。主要研究方向为机器学习, 小样本学习, 计算美学和风格迁移。

E-mail: luyue2016@ia.ac.cn

(LU Yue Ph.D. candidate at the State Key Laboratory for Management and

Control of Complex Systems, Institute of Automation, Chinese Academy of Sciences. His research interest covers



machine learning, few-shot learning, computational aesthetics, and style transfer.)



**郭超** 中国科学院自动化研究所复杂系统管理与控制国家重点实验室博士研究生. 主要研究方向为机器学习, 强化学习, 计算美学, 机器艺术创作和三维结构认知.

E-mail: guochao2014@ia.ac.cn

(**GUO Chao** Ph.D. candidate at the

State Key Laboratory for Management and Control of Complex Systems, Institute of Automation, Chinese Academy of Sciences. His research interest covers machine learning, reinforcement learning, computational aesthetics, machine art creation, and 3D structure cognition.)



**林懿伦** 中国科学院自动化研究所复杂系统管理与控制国家重点实验室助理研究员. 主要研究方向为社会计算, 智能交通系统, 深度学习和强化学习.

E-mail: yilun.lin@ia.ac.cn

(**LIN Yi-Lun** Assistant researcher at

the State Key Laboratory for Management and Control of Complex Systems, Institute of Automation, Chinese Academy of Sciences. His research interest covers social computing, intelligent transportation systems, deep learning, and reinforcement learning.)



**卓凡** 中央美术学院副教授. 主要研究方向为智慧(产品)设计, 传统造物方式与现代设计(美学)转型.

E-mail: zhuofan@cafa.edu.cn

(**ZHUO Fan** Associate professor at Central Academy of Fine Arts. His research interest covers intelligent

product design, aesthetic transformation of traditional creation method and modern design.)



**王飞跃** 中国科学院自动化研究所复杂系统管理与控制国家重点实验室主任, 中国科学院大学中国经济与社会安全研究中心主任, 青岛智能产业技术研究院院长. 主要研究方向为平行系统的方法与应用, 社会计算, 平行智能以及知识自动化. 本文通信作者.

E-mail: feiyue.wang@ia.ac.cn

(**WANG Fei-Yue** State specially appointed expert and director of the State Key Laboratory for Management and Control of Complex Systems, Institute of Automation, Chinese Academy of Sciences. Director of China Economic and Social Security Research Center in University of Chinese Academy of Sciences. Dean of Qingdao Academy of Intelligent Industries. His research interest covers methods and applications for parallel systems, social computing, parallel intelligence, and knowledge automation. Corresponding author of this paper.)