

中图法分类号: TP301.6 文献标识码: A 文章编号: 1006-8961(2018)08-1193-14

论文引用格式: Sheng J C, Li Y Z. Learning artistic objects for improved classification of Chinese paintings [J]. Journal of Image and Graphics, 2018, 23(8): 1193-1206. [盛家川, 李玉芝. 国画的艺术目标分割及深度学习与分类[J]. 中国图象图形学报, 2018, 23(8): 1193-1206.] [DOI: 10.11834/jig.170545]

国画的艺术目标分割及深度学习与分类

盛家川 李玉芝

天津财经大学理工学院, 天津 300222

摘要: 目的 现有对艺术画进行分类的文献大多对整幅画作直接进行特征提取, 但任何图像内容特征的可适应性都存在一定的局限性。画家画不同艺术目标的笔锋和艺术处理往往是不一样的, 如果不考虑每个笔锋所产生的条件而一味地分析笔锋的走向和力度分布等将会是很盲目的。为此提出一种基于艺术目标的中国画分类算法。方法 首先, 基于简单线性迭代聚类算法根据像素间颜色和位置的相差程度来生成超像素; 其次, 提出针对艺术目标的最大相似度区域合并算法交互式地进行艺术目标分割, 将国画分割成一系列的艺术目标, 如马、人物等, 以提取画家用来表现艺术形式和抒发情感的相对稳定单元; 然后利用深度卷积神经网络(DCNN)来描述这些艺术目标的高级语义特征; 最后, 引入支持向量机对每幅中国画内的各种艺术目标的分类结果进行最后的融合与分类。结果 本文针对艺术目标进行国画的学习和分类, 实现了对样本库中10位画家中国画的识别, 平均准确率为89%。实验结果表明, 本文算法在平均查全率和查准率上优于现有的MHMM(The 2D multi-resolution hidden Markov model)和Fusion等方法。结论 本文的成果可用于中国画的数字化分析、管理、理解和识别, 为中国画传承和鉴赏提供有效的数字工具。

关键词: 艺术目标分割; 中国画分类; 卷积神经网络; 融合算法; 深度学习; 超像素分割

Learning artistic objects for improved classification of Chinese paintings

Sheng Jiachuan, Li Yuzhi

School of Science and Technology, Tianjin University of Finance and Economics, Tianjin 300222, China

Abstract: **Objective** Presently, existing research on art classification is primarily based on feature extraction and hence feature-based classification. Although such feature-based methods reported in the literature achieve a certain level of success, a major weakness lies in the considerable dependence of classification performances on the effectiveness of the features in describing the content of Chinese paintings. Given that traditional Chinese artists tend to rely on popular objects, such as figures, trees, flowers, birds, mountains, horses, and houses to express their artistic feelings and emotions, we explore a new concept of artistic object-based approach to classify traditional Chinese paintings in this study. In this way, automated classification can be integrated with perception, understanding, and interpretation of artistic expressions and emotions via the segmented artistic objects. Such an approach also possibly enables our proposed methods to be further developed into an interactive object-based classification approach for other forms of paintings. In comparison with the existing state of the arts, one advantage of our proposed approach over those based on features or content is that objects provide direct and integrated art expressions inside paintings. **Method** Our proposed method includes three stages of processing and

收稿日期: 2017-10-25; 修回日期: 2018-03-07; 预印本日期: 2018-03-15

基金项目: 国家自然科学基金项目(61502331, 11701410)

Supported by: National Natural Science Foundation of China (61502331, 11701410)

analytics for traditional Chinese paintings ,that is ,1) interactive art object segmentation; 2) description and characterization of art objects via convolution neural network (CNN) ,the most popular deep learning unit; and 3) SVM-based classification and fusion across all art objects. Specifically ,via an iterative linear clustering algorithm ,super-pixels are constructed to detect the difference between the color and position of each individual pixel. By maximizing the similarity within the neighborhood of those super-pixels ,a sequence of objects can be segmented ,and an interactive scheme can be designed ,allowing users to add ,revise ,and interact with the content of paintings to achieve the best possible balance between subjective demand and objective art description. Afterward ,a CNN-based deep learning unit is added to describe those objects ,so its classification can be carried out with regard to the individual art object. Finally ,an SVM unit is adopted to achieve the final fusion of all these classifications via consideration of each individual object within the given window ,which is influenced and initialized through the training process. **Result** Extensive experiments are carried out ,which are in four phases ,each of which considers one impact factor ,such as consideration of the number of artists ,comparison with the existing state of the arts ,consideration of benchmarking via content-based classifications ,and assessment of contributions from CNN alone. Experimental results show that our proposed algorithm: 1) outperforms several existing representative approaches ,including MHMM and fusion-based method ,2) achieves effective fusion of all different object classifications ,including CNN and SVM units ,3) captures the artistic emotions through those segmented art objects ,and 4) shows potential for interactive classification of Chinese paintings via segmentation of artistic objects. **Conclusion** This study proposes the computerized classification and recognition of art styles based on artistic objects in paintings rather than the whole paintings. Experimental results reveal that the proposed algorithm outperforms the existing representative benchmarks ,providing potential for developing effective digital tools for computerized management of Chinese paintings. In addition ,this method can be used to formulate an important tool for computerized management of Chinese traditional paintings ,providing a range of techniques for effective and efficient digitization ,manipulation ,understanding ,perception ,and interpretation of Chinese traditional arts as well as its legacy.

Key words: artistic object segmentation; classification of Chinese paintings; convolutional neural network; fusion algorithm; deep learning; superpixel segmentation

0 引言

近年来 随着中国画数字化的飞速发展 如何建立和管理中国画数字图书馆和数字博物馆? 成为当前的一个热点研究问题 对中国画图像的处理技术是该研究问题的关键。如何发现这些画作背后的数字规律 如何有效地分析、识别和分类中国画的作者和艺术风格日趋成为研究热点。

目前大量文献利用计算机技术对艺术画进行分析和研究 但针对中国画的研究仍存在如下问题: 1) 现有艺术画方面的成果主要是将自然图像进行油画、粉笔画等艺术风格模拟 或者是针对油画等西方画进行艺术风格分析^[1]。国画作为一种独特的艺术形式 与其他艺术风格(如油画、卡通、粉笔画等) 相比有着截然不同的意境和韵味 因此一些已有的针对西方画等的研究不能直接用于中国画的分析中。2) 目前针对中国画的科研工作较少 主要研究工作如下: 盛家川和江健民^[2]提取了整体和局部

特征 并提出熵平衡(Fusion) 算法对国画作者进行分类研究。盛家川^[3]研究了在小波域内不同尺度和频带上画作所体现出的不同深度信息特征 以对画作进行分类。Li 等人^[4]设计了相关算法分类沈周、唐寅、张大千等中国画家的画作 该算法首先对画作提取小波特征 并设计了 MHMM 隐马尔可夫模型对画家进行分类研究。Jiang 等人^[5]提出了一种把中国画分为写意和工笔两类的算法 首先提取颜色、纹理等底层特征 进而采用支持向量机进行分类。王征和孙美君^[6-7]提出了复合特征的自适应遴选及对水墨画风格描述的优化 其还通过提取多种底层异构视觉特征对国画风格进行描述 对国画作者进行分类预测。Fan 等人^[8]提出了一种色彩转换算法将照片转换成中国水墨画。目前 针对中国画的工作大多集中在基于内容的分析和检索(content-based image analysis) 但存在以下问题: 1) 任何图像内容特征的可适应性都存在一定的局限性。例如 画马时的笔锋与画树叶时的笔锋肯定是不同的^[9-10] 如果不考虑每个笔锋所产生的条件而一味地分析笔锋

的走向和力度分布等将会是很盲目的。2) 现有基于内容的研究会考虑画作中所有的信息, 这样会使得研究重点较为分散, 容易受到噪声信息的干扰。

针对以上问题, 本文将国画中具有艺术风格的绘画对象定义为“艺术目标(artistic object)”, 如国画中的花、鸟、人物、树等, 这些艺术目标是画家用来表现艺术形式和抒发情感的相对稳定单元, 是画作艺术风格特征表达的载体。进而, 本文提出一个对中国画中的主要艺术目标进行交互式分割及识别的框架, 以便针对艺术目标进行数字量化、分析, 提炼出最能体现画家艺术风格的高级语义信息。

首先, 基于简单线性迭代聚类(SLIC)算法根据像素间颜色和位置的相差程度进行超像素网格分割; 第二, 提出最大相似度区域融合算法(MSRMAO)交互式地进行艺术目标分割, 即将整幅画作分割成一系列的艺术目标, 如马、人物等, 以提取画家用来表现艺术形式和抒发情感的相对稳定单元; 第三, 基于深度学习提取艺术目标的特征描述(O-CNN), 以提取这些艺术目标的数字特征; 最后, 提出一种基于支持向量机的融合算法对提取出的艺术目标进行学习和识别, 进而实现基于艺术目标的中国画艺术风格分析和识别。

基于艺术目标的国画分类识别体系具有以下优势: 1) 避免盲目的对整幅国画内容进行特征提取, 而是将最能代表画家风格的艺术目标作为研究重点, 从而可以更加细粒度地对国画风格进行分析。并且将画家的艺术风格分析集中到局部的一些主流艺术目标上, 使研究更加集中有效, 从而更好的对画家和艺术风格进行理解。2) 利用CNN在特征提取方面的巨大优势针对艺术目标提取数字特征, 这样避免了显示的特征提取, 转而在训练中自动隐式地学习; 并且因为同一特征映射面上的神经元权值相同, 从而网络能够并行地学习。3) 对于一幅中国画中含有多个艺术目标的情况, 本文的交互式分割设计可以更加贴合鉴赏者的主观意愿, 以最大程度上实现主观需求与机器反馈的互动交流, 也建立了一个新的角度去研究画家在创作时对每个艺术目标的考量和艺术风格运用。

1 相关算法

超像素^[11], 是指图像中像素点组成的子区域,

属于同一区域的像素点的性质相差较小, 比如颜色、空间位置相近。超像素分割算法对图像进行分割即可生成超像素, 这些超像素能保留图像的边界结构, 大大减少了冗余像素, 这样可使图像后续处理变得更容易。

近年来有两种常用的产生超像素的方法: 基于图论的算法和基于梯度上升的算法。前者是构造的能量函数最小化问题, 将图像构造成一个带权无向图, 图像中像素点对应图节点, 两个像素的相邻关系表示图的一条边, 而边的权值是指相邻像素的相差程度。对图划分时以划分后子图的局部相似度最大为原则, 从而生成超像素。包括 SuperLattice^[12]、EGS^[13]、Ncut^[14]等算法。后者的基本思想是从初始的种子点开始, 每一次迭代利用一定准则对像素进行聚类, 直到达到稳定状态, 从而生成超像素, 包括 MeanShift^[15]、TurboPixel^[16]和 SLIC^[17](simple linear iterative clustering)等算法。

SuperLattice^[12]算法使用贪婪策略, 每次在边界成本图最小处利用水平和垂直路径来分割图像, 从而得到超像素。该方法保持了规则的图像拓扑结构, 产生规整的超像素网格, 具有良好的分割精度和稳定性, 同时超像素数量是可以人为规定的。但该方法产生的超像素优劣与图像的边界图质量有很大关系。

MeanShift^[15]算法是一种无参数的迭代算法, 通过概率密度函数使中心点收敛至密度最大的点。该方法产生的超像素形状规整, 在稳定性和抗躁性上保持良好的性能。但是该方法速度不快, 对于超像素数量无法控制, 并且存在过分割问题。

TurboPixel^[16]算法是基于几何流的水平集方法, 首先选择初始种子点, 通过曲率演化模型和骨架化过程来扩张种子点的区域, 从而得到网格状超像素。该算法的运行时间与图像尺寸是正相关关系, 能够人为规定生成的超像素数目, 超像素形状规则且能保留图像的轮廓结构, 同时改良了欠分割问题。但是生成的超像素形状不可控, 对于分辨率较大的图像, 不能满足快速高质量的图像分割。

SLIC^[17]算法是一种基于聚类算法的超像素分割, 由像素LAB颜色空间和像素位置共5维空间来生成超像素。该方法首先初始化聚类中心, 并通过重设聚类中心将其移到邻域内梯度最小的地方, 然后在聚类中心 $2S \times 2S$ 的邻域内, 为每个聚类中心分

匹配点,进而计算新的聚类中心与之前聚类中心的距离,根据阈值判断是否需要重新设置聚类中心。该方法生成的超像素大小均匀、形状规整,同时边界信息保持较好,能控制生成的超像素的数目,且该数量是算法的唯一输入参数。

对于本文的艺术目标分割问题,我们需要构建一个规整的超像素网格。SLIC 算法和 MeanShift、TurboPixel 均可产生规整超像素,而 SuperLattice 分割质量受输入边界图的影响,MeanShift 不能控制生成的超像素数量,TurboPixel 对于分辨率较大的图像性能较差。综合比较,SLIC 算法的运行速度比目前超像素分割算法快,同时生成的超像素形状规整且保持图像局部结构,因此,本文的图像分割阶段使用 SLIC 算法。

2 国画艺术目标的深度学习、融合与分类算法

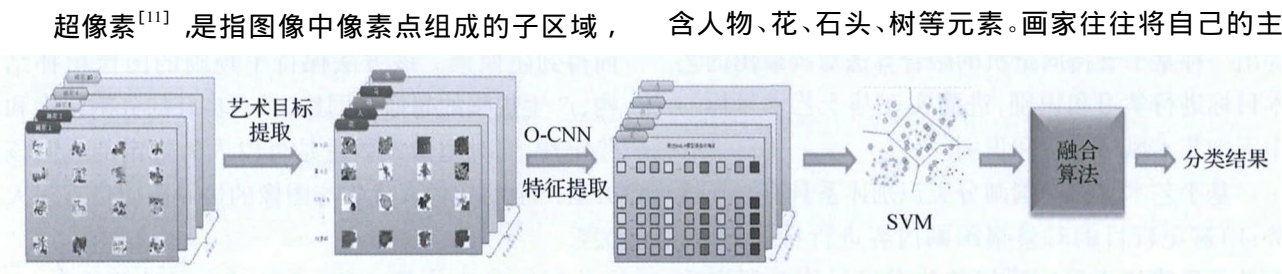


图1 国画艺术目标的深度学习、融合与分类算法整体示意图
Fig.1 Illustration of our system

观情感、思想等通过画中的艺术目标进行表现,可以说这些艺术目标在一定程度上成为画家思想感情表达的一种载体^[10,18]。由于中国画艺术目标往往是有限的,并且每位画家的手法、情感、习惯、性格等存在很大差异,这就造成不同的画家对于同一种艺术目标往往赋予不同的理解和感情,使得这些艺术目标各具特色^[18]。

如图2所示,同样的艺术目标荷花,不同画家赋予了它不同的性格、艺术理解和思想感情,蕴藏着迥异的数字艺术特征。齐白石笔下的荷图大多拥有固有的乐观健朗的神气,无论是秋荷、残荷、还是缀以虫鸟的荷趣图,荷花的形象皆苍劲而清健,挺拔而厚重。八大山人笔下的荷花不是春色娇艳、欣欣向荣的鲜花,而是“溅泪”之墨花,是残叶败荷,一如他一向的画风,凄凉寂寞,冷意逼人。潘天寿的荷花笔墨

属于同一区域的像素点的性质相差较小,比如颜色、空间位置相近。超像素分割算法对图像进行分割即可生成超像素,这些超像素能保留图像的边界结构,大大减少了冗余像素,这样可使图像后续处理变得更容易。

本文提出的国画艺术目标的深度学习、融合与分类算法由3个模块组成。第1个模块提出一种针对艺术目标的最大相似度区域融合算法MSRMAO,用以从中国画中交互式地提取艺术目标。第2个模块构造了一个针对艺术目标的卷积神经网络(O-CNN),用以对每个艺术目标提取一个定长的数字特征向量。第3个模块是艺术目标分类结果的融合。本文算法整体示意图如图1所示。

2.1 中国画的艺术目标研究

中国画所描绘的艺术目标往往是有限的^[9-10],以国画中的“山水画”为例,其内容大多都包含山、石、树、房子等元素,“花鸟画”其内容大多都包含花、鸟、树、石头等元素,“人物画”其内容大多都包含人物、花、石头、树等元素。画家往往将自己的主

苍古、凝炼老辣,且大气磅礴,雄浑奇崛,具有摄人心魄的力量感和现代结构美。张大千的荷花素净纯美,亭亭玉立又不失潇洒飘逸,将“静若处子,动若脱兔”两种对立之美糅合其中,远观之仙气十足,与周敦颐莲之“不可亵玩”的高洁性情甚符。大多数中国画作品,其内容大多都包含花、鸟、人物等元素,并且画家会将自己的主观情志凝结于这些艺术目标之上,从而导致不同画家笔下的相同艺术目标有不同的表达方式。因此,针对艺术目标对中国画进行分析是非常有必要的。

另一方面,任何图像内容特征的可适应性都存在一定的局限性,因此针对艺术目标进行研究,将优于对整幅图像直接进行特征提取。例如,画马时的笔锋与画树叶时的笔锋肯定是不同的^[10],如果不考虑每个笔锋所产生的条件而一味地分析笔锋的走向



图2 不同画家笔下性格迥异的荷花

Fig. 2 Study upon artistic styles of lotuses comparison across different artists ((a) Qi Baishi; (b) Bada Shanren; (c) Pan Tianshou; (d) Zhang Daqian)

和力度分布等将会是很盲目的。因此,如何将国画分割成一系列的艺术目标,如马、人物等,以提取画家用来表现艺术形式和抒发情感的相对稳定单元,进而实现基于艺术目标的风格分析非常重要。

因此,本文提出了一个国画交互式艺术目标分割的框架,利用该框架可以有效提取出画作中相对稳定的艺术目标,从而基于这些艺术目标提取画作的风格特征,更好地为分类识别奠定基础。

2.2 交互式艺术目标分割

本文研究如何将国画交互式分割成一系列的艺术目标,如花、鸟、人物等,以提取画家用来表现艺术形式和抒发情感的相对稳定单元。

设给定的国画库中包括 M 个传统艺术家,每个艺术家的收藏作品数用 $\{N_1, N_2, \dots, N_M\}$ 表示,则国

画库的数字化描述可由图3所示。因此对第 j 个艺术家,其国画艺术目标的计算机描述可表示为 $\{I_j(G_i) | j \in [1, M]; i \in [1, N_j]\}$ 。 G_i 表示该艺术家 N_j 个作品中第 i 幅国画,将 G_i 分割成一系列的主要艺术目标 $G_i = \{O_1^i, O_2^i, \dots, O_k^i\}$,式中每一个 O_k^i 都代表通常在一幅国画作品中见到的艺术目标,例如花、鸟和人物等,这些有限的艺术目标是艺术表达的基本载体。

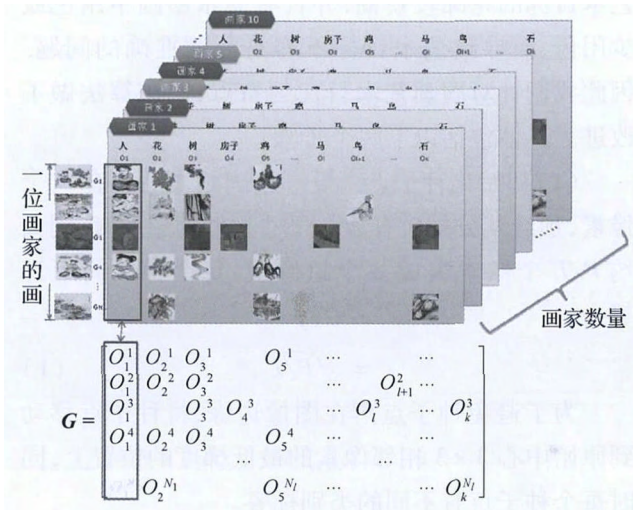


图3 基于艺术目标分割的国画数据库数字化描述示意图

Fig. 3 Dataset of the artistic objects

为了将 G_i 分割成一系列的主要艺术目标 $G_i = \{O_1^i, O_2^i, \dots, O_k^i\}$,本文提出了一个交互式的国画艺术目标分割方案,如图4所示。该方案由两部分组成,第1部分基于 SLIC^[17] 生成超像素网格,第2部分提出一种最大相似度区域融合的中国画艺术目标提取算法 (MSRMAO),用户只需要输入一些交互信息,概念性的用简单划线指示出艺术目标和背景,算法将自动完成艺术目标的提取工作。

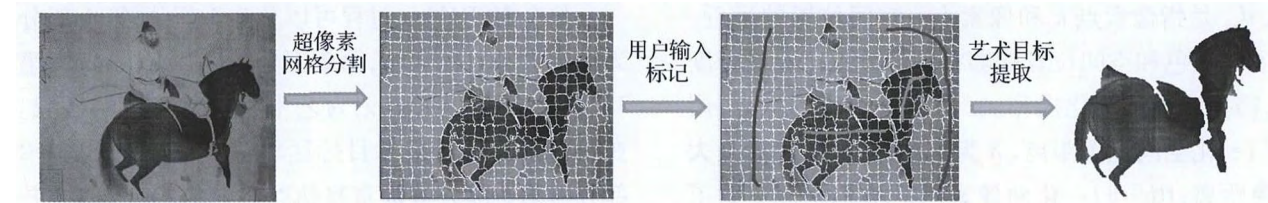


图4 国画交互式艺术目标分割方案图

Fig. 4 Illustration of interactive artistic object segmentation

艺术目标分割方案的两部分内容如下:

1) SLIC 生成超像素网格。本文基于 SLIC^[17] 算法对中国画进行超像素网格分割,将具有相似纹理、

颜色、亮度等特征的相邻像素结合成许多不规则的像素区域。它利用像素之间特征的相似性将像素分组,用少量的超像素代替大量的像素来表达图片特

征,很大程度上降低了提取中国画艺术目标的复杂度。

SLIC^[17]的思想是首先用 5 维特征向量表示像素点的特征,该特征向量包括 CIELAB 颜色空间和空间坐标 XY,然后对图像像素进行局部聚类从而得到超像素。该算法运行时间较少,且可以产生规整的超像素,但应用在国画的艺术目标分割时,由于中国画具有独特的运笔用墨技巧,与自然图像相比,艺术目标的轮廓较模糊,并且有的水墨画不用色或少用色,而导致艺术目标不容易分割准确的问题。因此我们针对国画艺术目标的特点,对该算法做了改进。具体包括以下两个步骤:

(1) 初始化种子点。设一幅国画作品共有 P 个像素,预计其分割后有 L 个尺寸相似的超像素,则大约 P/L 个像素组成一个超像素,每个种子间距计算为

$$s = \sqrt{P/L} \quad (1)$$

为了避免种子点落在图像边缘,将种子点移动到原始中心 3×3 相邻像素的最低梯度的位置上,同时每个种子点有不同的类别标签。

(2) 权衡相似程度。分别计算每个像素点与它距离最近的种子点的相差程度,若像素点与种子点 P 差异最小,则该像素点与种子点 P 的类别标签相同。重复这个过程,直到达到稳定状态。其相关的描述为

$$d_{lab} = \sqrt{(l_k - l_i)^2 + (a_k - a_i)^2 + (b_k - b_i)^2} \quad (2)$$

$$d_{xy} = \sqrt{(x_k - x_i)^2 + (y_k - y_i)^2} \quad (3)$$

$$D = \sqrt{\left(\frac{d_{lab}}{\delta}\right)^2 + m\left(\frac{d_{xy}}{s}\right)^2} \quad (4)$$

式中, d_{lab} 是指像素点 k 和像素点 i 颜色的相差程度, d_{xy} 是指像素点 k 和像素点 i 空间位置的远近。为了使颜色和空间位置能够在同一个尺度上进行度量,需要进行归一化操作。 s 为种子点间的距离,用于归一化空间位置距离。 δ 为 LAB 色彩空间中最大颜色距离,用于归一化两像素间的颜色距离。为了综合考虑中国画水墨颜色与空间位置的重要程度,本文引入权重变量 m ,以衡量像素点间的相似度。 D 值代表了像素点间的差异程度, D 值越大,表示像素点间差异越大。

为种子点划分像素时,没有在整张图像中寻找,而是在以种子点为中心的 $2S \times 2S$ 区域内搜索像素

点,这能够极大加快算法运行速度。

2) MSRMAO 算法交互式提取艺术目标。与自然图像不同,中国画属于艺术作品,其中包含的艺术目标具有一定的概念性和抽象性,尤其是中国画中的泼墨手法,一般画作中的艺术目标没有明确的轮廓勾勒。因此,本文提出一种最大相似度区域融合的中国画艺术目标提取算法 MSRMAO。MSRMAO 算法通过用户交互式地给出艺术目标与背景区域的笔道信息,基于最大相似度区域合并规则,可以完成区域合并,提取艺术目标。

将标记引入交互式分割的想法在 Meyer 的 watershed 方案^[19]中已被提出,但本文用它指导区域合并来对中国画艺术目标的轮廓进行提取。该算法优点是能够在复杂背景下提取出整幅中国画中的艺术目标,无需预先的阈值和参数设置,可以获取较好的分割效果。

MSRMAO 算法利用颜色直方图刻画图像特征,将 RGB 颜色空间的每个颜色通道量化成 16 个等级的颜色亮度值,那么每块区域的颜色直方图都可以在一个 $16 \times 16 \times 16 = 4096$ 组的颜色空间中来计算某个像素点的亮度值。用 H_S 、 H_T 分别表示区域 S 和 T 的归一化直方图。在交互式分割中,用户所标记的艺术目标或背景区域和未标记的区域需要基于一定的相似性规则进行合并。这里,选择采用巴氏系数(Bhattacharyya) ζ 来测量区域 S 和 T 之间的相似度,计算公式为

$$\zeta(S, T) = \sum_{\eta=1}^{4096} \sqrt{H_S^\eta \times H_T^\eta} \quad (5)$$

将通过用户标记需要提取的艺术目标区域的集合,记为 F ; 背景区域的集合,记为 B ; 中国画中剩余未标记的区域集合,记为 M 。

整个 MSRMAO 过程可以分为 3 部分,第 1 部分为融合艺术目标区域,第 2 部分为融合背景区域,第 3 部分为融合仍未标记为艺术目标或背景的区域。在第 1 部分融合艺术目标区域执行结束之后,第 2 部分与第 3 部分将被重复执行,直到没有新的合并发生。本文策略是在粗略融合艺术目标区域后,更多的合并所有的背景区域,直到所有的区域都被融合。

MSRMAO 算法步骤如下:

输入: 被用户用笔道标记了前背景的原图以及利用 SLIC 算法生成的超像素网格。

输出: 最终迭代完成分割出的艺术目标图像。

算法步骤:

1) 对中国画艺术目标区域 F 进行融合:

(1) 对每个区域 $P \in F$, 标记其相邻区域为 $\bar{R}_P = \{A_i\} \quad i=1, 2, \dots, p$ 。

(2) 对每个 $A_i \notin F$, 标记其相邻区域为 $\bar{R}_{A_i} = \{R_j^{A_i}\} \quad j=1, 2, \dots, k$ 。显然有: $P \in \bar{R}_{A_i}$ 。

(3) 计算 $\zeta(A_i, R_j^{A_i})$, 如果 $\zeta(A_i, P) = \max_{j=1, 2, \dots, k} \zeta(A_i, R_j^{A_i})$, 则将 P 和 A_i 合并 $P = P \cup A_i$, 否则不合并。

(4) 更新 F 和 M 。

(5) 如果 F 中未找到新的可合并区域, 则进行步骤 2); 否则返回步骤 (1)。

2) 对背景区域 B 进行融合:

(1) 对于每个区域 $Q \in B$, 标记其相邻区域为 $\bar{R}_Q = \{N_i\} \quad i=1, 2, \dots, q$ 。

(2) 对每个 $N_i \notin B$, 标记其相邻区域为 $\bar{R}_{N_i} = \{R_j^{N_i}\} \quad j=1, 2, \dots, k$ 。显然有: $Q \in \bar{R}_{N_i}$ 。

(3) 计算 $\zeta(N_i, R_j^{N_i})$, 如果 $\zeta(N_i, Q) = \max_{j=1, 2, \dots, k} \zeta(N_i, R_j^{N_i})$, 就将 Q 跟 N_i 合并 $Q = Q \cup N_i$, 否则不合并。

(4) 更新 B 和 M 。

(5) 如果 B 中找不到新的可合并区域, 则进行下一步; 否则返回步骤 (1)。

3) 对未标记区域 M 进行融合:

(1) 对于每个未标记区域 $O \in M$, 标记其相邻区域为 $\bar{R}_O = \{G_i\} \quad i=1, 2, \dots, \rho$ 。

(2) 对于每一个 $G_i \notin B$ 且 $G_i \notin F$, 标记其相邻区域 $\bar{R}_{G_i} = \{R_j^{G_i}\} \quad j=1, 2, \dots, k$ 。显然 $O \in \bar{R}_{G_i}$ 。

(3) 计算 $\zeta(G_i, R_j^{G_i})$, 如果 $\zeta(O, G_i) = \max_{j=1, 2, \dots, k} \zeta(G_i, R_j^{G_i})$, 合并 O 与 G_i $O = O \cup G_i$, 否则不合并。

(4) 更新 M 。

(5) 如果 M 中没有新的可合并区域, 则进行下一步; 否则返回步骤 (1)。

4) 返回 2) 进行循环, 直到没有可融合的区域, 结束融合, 除背景区域外都会被提取为艺术目标。

如图 4 所示, 欲提取的前景艺术目标用绿色笔道简单标识出来, 背景用蓝色笔道画出, 通过几次迭代算法最终艺术目标完成融合并成功与背景分离,

取得了良好的分割结果, 算法在时间效率上也相对较好。

从分割效果图 4 可以看出, 由于 SLIC 分割较好地保留了艺术目标的边缘信息, 所以合并后的艺术目标与背景的分离程度很好。通过大量实验结果对比, 发现每个画家笔下的艺术目标风格具有很高的稳定性和相似性, 然而不同画家之间对于相同艺术目标的描绘又存在着视觉上的差异, 如图 8。本文将代表画家绘画艺术风格特点且情感表达最稳定的艺术目标按照不同类别分类提取, 通过局部特征的分析对比, 将某个画家的艺术风格集中到局部的艺术目标上, 从而为后续特征提取与分类建立良好的基础。

另外, 对于一幅中国画中含有多个艺术目标的情况, 本文的互动式设计可以让用户添加、修改并与绘画内容互动, 从而更加贴合鉴赏者或是初学画者的主观意愿。他们可以根据自己对整幅作品中最感兴趣的部分进行重点研究, 而不是把整幅画作作为研究对象, 以达到主观需求与客观艺术描述之间的最佳平衡。本文对于艺术目标的分割提取, 是为了建立一个新的角度去研究画家在创作时对每个艺术目标的考量和艺术风格运用。后续实验也将表明每位画家笔下的艺术目标的确具有稳定性和鲁棒性。

2.3 基于深度学习的艺术目标特征描述 (O-CNN)

深度学习模仿人脑机制, 组合低层特征, 获取表示数据分布式特征的高层特征, 并广泛应用于手写识别、人类行为识别、目标检测等方面^[20]。近年来, 深度学习也被用于图像细粒度分析与分类的研究并取得了很好的结果^[21-24]。

传统的卷积神经网络 (CNN) 一般直接用于图像分类。本文利用 CNN 在特征提取方面的巨大优势对中国画艺术目标进行特征描述, 并将其定义为 O-CNN 算法, 相比传统显示的特征提取方法, O-CNN 能并行地自动学习训练数据中的低层次信息, 得到更详细高层的特征数据, 从而对每一个艺术目标都提取一个 4096 维的定长特征向量。算法首先将分割出的艺术目标转换为 224×224 像素的图片, 作为模型的输入。其次, 基于 CNN 理论框架, 通过微调 ImageNet 数据集上预训练的 VGG-F 模型^[25]来提取中国画稳定艺术目标的图像特征, 该模型是一种前馈神经网络。本文利用 VGG-F 模型的 5 个卷积层和 2 个全连接层进行特征提取, 并将第 2 个

全连接层作为艺术目标的特征进行输出,如图5所示。

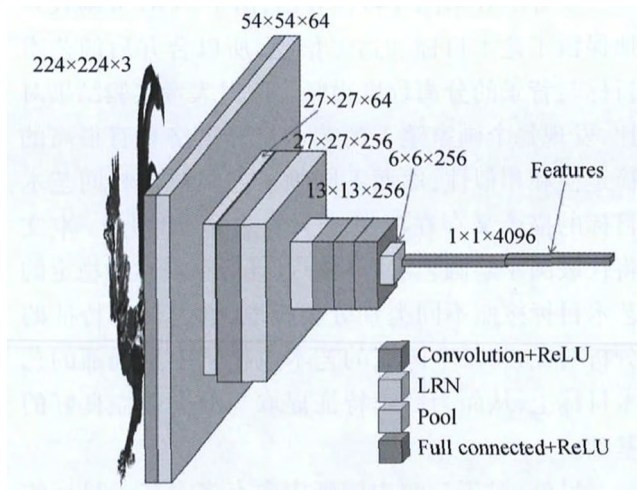


图5 CNN特征提取示意图

Fig. 5 Structure of CNN based feature extraction

利用自然图像预训练的 VGG-F 模型对中国画进行特征提取是可行的。首先,艺术来源于生活,画家通过对自然万物的分析和体悟来完成作品,即中国画是对自然场景的艺术再加工,是自然图像的另一种体现。其次,VGG-F 模型的深度结构可以从丰富的感知信息中提取复杂的结构和建立数据中内在的表征。参与训练的自然图片有一千多万张,提取的特征会直接或间接包含和中国画相类似的特征信息,从而能学习到用于中国画特征提取的有用信息。最后,虽然本文训练数据集中的中国画艺术目标数量不足以完成 VGG-F 模型的训练,但是模型已在 ImageNet 大量图片下进行了充分的预训练,并且本文通过将中国画艺术目标训练数据集中的图像进行变换(缩放、旋转、裁剪、加噪声等方法),增加了一定的样本量,并将艺术目标作为引导加入到 VGG-F 模型中,对模型进行了微调,使得模型提取的特征更能表述中国画艺术风格。

2.4 艺术目标分类结果的融合

本文基于 LIBSVM 工具箱对艺术目标进行画家分类,该工具箱的多分类是基于一对一算法(one-versus-one),其多分类策略是在任意两类样本之间设计一个 SVM,因此 k 个类别的样本就需要设计 $k(k-1)/2$ 个 SVM,利用该策略对一个未知样本进行分类时,最后得票最多的类别即为该未知样本的类别。本文使用 LIBSVM 工具箱进行分类,通过穷

举法对非线性核函数及其惩罚系数以及其他参数进行调整,以得到更优化的分类结果。

一幅画往往会提取出多个艺术目标,依照本文算法对艺术目标进行学习和识别后,可能存在一幅画作中的多个艺术目标被识别成不同画家的情况。在这种情形下,就需要针对不同艺术目标的识别结果进行融合。有关融合的研究虽然历史悠久,已发表的融合方法很多,但针对不同艺术目标的分类结果存在差异的情况下获得最有效的融合仍然存在较大的挑战。鉴于此提出如下的融合方案。针对一系列的艺术目标 A, B, \dots, N 和所获得的分类结果 S_A, S_B, \dots, S_N , 最后的融合公式为

$$S_f = \arg \max_{W \subseteq \{A, B, \dots, N\}} \{P(S_i), \forall i \in W\} \quad (6)$$

式中, W 是某一幅画中所有艺术目标的子集,也可以看做是一个分类实验的窗口,主要基于平衡全局和局部最优的考虑而设立的控制分类结果正确概率的范围,该窗口 W 的大小决定了参与融合的艺术目标个数。一方面窗口设计过大会噪声进来,局部干扰增加,从而影响分类准确度;另一方面窗口设计过小,分类结果容易产生波动而不稳定,并且由于没有包含足够的信息,容易忽略局部艺术风格变化的影响,从而导致分类的结果产生偏差。因此,窗口过大或者过小都会导致分类结果下降。本文实验中,由于实验数据中的艺术目标不是很多,窗口 W 的大小取值为 12,以尽量多的包含所有艺术目标,并且本文取最高概率对应的画家类别,该策略也在一定程度上规避了噪声对分类结果的影响。 S_i 为艺术目标 i 的分类结果, $i \in W$, $P(S_i)$ 为当前该艺术目标 i 分类结果正确的概率,其初始值由训练结果获得。

本文提出的基于 SVM 融合算法在实现过程中也取得了反应速度快,计算代价低的优点。具体来说,首先对已提取出的九类艺术目标:房子、花、鸡、马、鸟、人、山石、树、鱼分别进行画家的分类测试。实验结果会按当前窗口的控制统计基于每个艺术目标对画家分类正确的概率,然后取最高概率对应的画家类别。具体的流程如图6所示。

3 实验结果及分析

为了检测本文算法的有效性,建立了中国画数据集。该数据集的建立得到了天津美术学院的帮

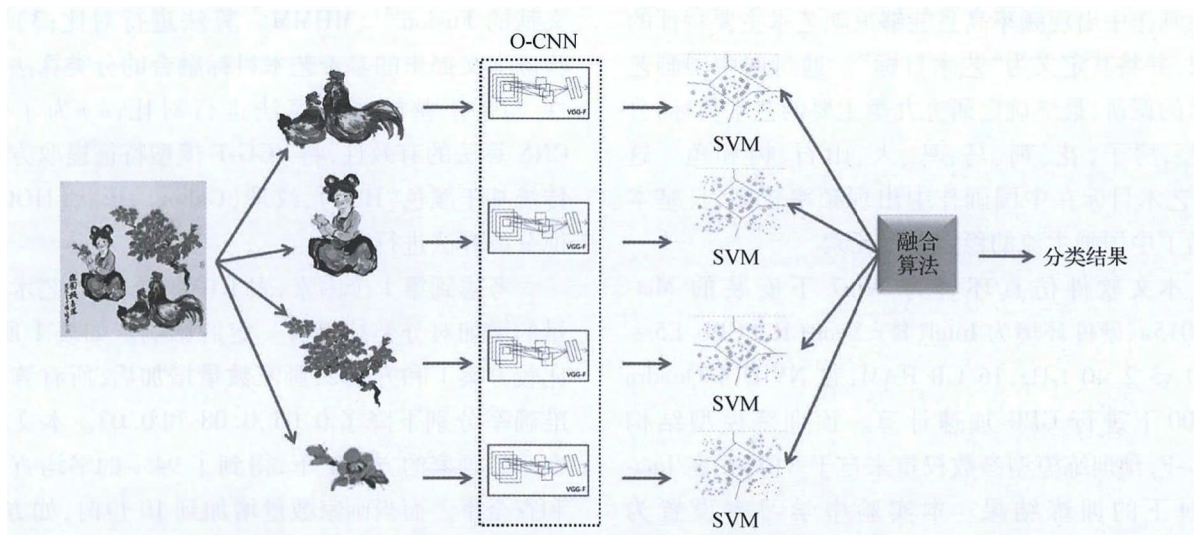


图6 艺术目标分类结果的融合

Fig. 6 Fusion of the classifications via consideration of each individual object

助,并从博物馆、雅昌艺术网、Flickr 网络下载、画家个人画集等途径,收集了10位中国画画家各100幅画作,并将其中的70%画作作为训练样本集,剩余部分作为测试样本集,训练与测试样本集中的数据不重叠。

在构建中国画数据集选取画家时,考虑到了以下两方面的因素:1)考虑到验证本文算法的有效性,收录了一些擅长绘画各种艺术目标的画家,即该画家创作的作品中包含4-5个艺术目标,以考查对艺术目标识别融合算法的有效性。例如,近当代国画界的“开门人”刘奎龄大师,就以“全能画家”著称,他擅长各类山水、动物以及人物画。2)考虑到中国画画家所处年代的不同,数据集中收录的画家从古代到当代在时代上有很大的跨越。考虑到以上两方面的因素,构建的中国画数据集中包含的画家分别是:陈大羽(1912—2001年)、程十发(1921—2007年)、范曾(1938年—)、黄永玉(1924年—)、黄胄(1925—1997年)、刘旦宅(1931—2011年)、刘奎龄(1885—1967年)、唐寅(1470—1524年)、徐悲鸿(1895—1953年)、八大山人(1626—1705年),这其中既包括艺术风格独特有见地的大师,也有当代虽未耳熟能详但艺术风格独树一帜的画家,这为后续通过特征提取分析画家的艺术风格奠定了良好的基础。从数据集中随机挑选的部分画作如图7所示,从上到下每一行是一位画家的画作,他们分别是陈大羽、程十发、范曾、黄永玉、黄胄、刘旦宅、刘奎龄、唐寅、徐悲鸿、八大山人。



图7 国画图片库中具有代表性的作品

Fig. 7 Representative samples of the data set for experimental purposes

通过调研,发现中国画的艺术目标是有限的,同一艺术目标会在不同画家的画作中反复出现^[9-10,18]。国画中物体种类繁多,本文着重研究中

国画画作中出现频率高且能够反映艺术主要特征的对象,并将其定义为“艺术目标”。通过向中国画艺术家的调研,最终确定研究九类主要的艺术目标,分别是:房子、花、鸡、马、鸟、人、山石、树和鱼。这9类艺术目标在中国画作中出现频率较高,且基本囊括了中国画主流的稳定艺术元素。

本文软件仿真环境为 Win7 下安装的 Matlab2015a,硬件环境为 Intel(R) Xeon(R) CPU E5-2630 v3 2.40 GHz,16 GB RAM,在 NVIDIA Quadro K1200 下进行 GPU 加速计算。预训练模型结构 VGG-F,预训练模型参数权重来自于 VGG-F 在 ImageNet 下的训练结果。本实验中学习率设置为 0.001,迭代次数设为 1 000,批训练大小为 50。

本文采用查准率和查全率^[1]来评价算法的结果。为了检测本文算法,从以下 4 个方面考虑设计实验:1)为了验证画家数量的变化对分类准确度的影响,分别对 5 位画家、10 位画家的作品集合进行分类实验对比;2)为了验证算法的有效性,与现有

文献的 Fusion^[1]、MHMM^[3] 算法进行对比;3)为了验证本文提出的基于艺术目标融合的分类算法有效性,与基于整幅画作算法进行对比;4)为了验证 CNN 算法的有效性,将 VGG-F 模型特征提取方法与传统基于颜色(HSV)、纹理(Gabor)、形状(HOG)特征分类算法进行对比。

考虑到第 1 个因素,表 1 中数据表明,艺术家数量的增加对分类结果有一定的影响。如表 1 所示,比较方案 1 和方案 2,画家数量增加后,所有算法的准确率分别下降了 0.05、0.08 和 0.07。本文算法在 5 位画家的方案 1 下,得到了 94% 的平均查准率和查全率。而当画家数量增加到 10 位时,如方案 2 所示,本文算法虽有所下降,但仍然得到了 89% 的平均查准率和查全率。并且相比 Fusion 算法和 MHMM 算法,本文算法下降幅度最小,分类准确率最高。实验结果表明,当画家数量增加时,本文算法仍然具有较高的分类准确率,能够根据画家不同风格进行识别具有一定的鲁棒性。

表 1 本文算法与 Fusion、MHMM 算法实验结果对比
Table 1 Experimental results compared with Fusion and MHMM algorithm

案例	画家	本文算法		Fusion 算法		MHMM 算法	
		<i>P</i>	<i>R</i>	<i>P</i>	<i>R</i>	<i>P</i>	<i>R</i>
1	陈大羽	0.93	0.93	0.83	0.80	0.79	0.77
	程十发	0.97	0.97	0.81	0.83	0.83	0.80
	刘奎龄	0.96	0.90	0.79	0.77	0.84	0.90
	徐悲鸿	0.88	0.97	0.84	0.90	0.84	0.87
	八大山人	0.97	0.93	0.90	0.87	0.90	0.87
	平均值	0.94	0.94	0.83	0.83	0.84	0.84
2	陈大羽	0.90	0.93	0.76	0.73	0.75	0.70
	程十发	0.94	0.97	0.74	0.77	0.73	0.73
	范增	0.85	0.93	0.72	0.77	0.80	0.80
	黄永玉	0.93	0.90	0.79	0.77	0.74	0.77
	黄胄	0.79	0.73	0.59	0.53	0.62	0.60
	刘旦宅	0.81	0.87	0.72	0.70	0.74	0.77
	刘奎龄	0.92	0.80	0.73	0.73	0.81	0.87
	唐寅	0.93	0.87	0.78	0.83	0.89	0.83
	徐悲鸿	0.88	0.97	0.81	0.87	0.81	0.83
	八大山人	0.93	0.90	0.86	0.83	0.83	0.83
	平均值	0.89	0.89	0.75	0.75	0.77	0.77

考虑到第 2 个因素,从表 1 中可以看到,本文算法与 Fusion 算法和 MHMM 算法比较都有最高的分类准确率,表明该算法可以更有效地反映画家不同的艺术风格,从而得到更好的画作作者分类结果。

考虑到第 3 个因素,如表 2 所示,将本文算法与整幅图通过 CNN 提取特征的分类结果作对比。实验结果表明,基于艺术目标的特征描述比基于整幅图像更有效,说明针对艺术目标的描述优于整体画作的描述。实验结果表明,相比整幅图,除徐悲鸿没有提高外,其余画家在艺术目标特征的描述符下都显示出了更好的识别结果。因为徐悲鸿的画作中艺术目标比较单一,往往一幅画作中只有一匹马,即只包含一个艺术目标,这样就导致本文的艺术目标融

合算法无法发挥其优势作用。而本文算法对于一幅图像中存在多个艺术目标的作品分类正确率会有很大的提高,如陈大羽和程十发两位画家画作中艺术目标丰富,应用本文算法后,其查全率分别从 87% 和 90%,提高到 93% 和 97%。

考虑第 4 个因素,为了验证本文 CNN 算法提取特征的有效性,将 VGG-F 模型特征提取方法与通过传统方式提取特征的分类结果作对比,其中传统特征提取包括基于颜色(HSV)、纹理(Gabor)、形状(HOG)的特征,如表 2 所示。实验结果表明,本文 VGG-F 模型特征提取算法相比传统算法,得到了更高的准确率,证明了 VGG-F 模型提取的特征能更有效地反映画家的不同艺术风格。

表 2 本文算法与整幅图像和传统特征提取算法的实验结果对比
Table 2 Comparative experimental results between the proposed, the whole image based methods and hand-crafted features methods

画家	本文算法		整幅图通过 CNN 提取特征		整幅图通过 HSV + Gabor + HOG 提取特征	
	P	R	P	R	P	R
陈大羽	0.93	0.93	0.90	0.87	0.72	0.70
程十发	0.97	0.97	0.90	0.90	0.59	0.63
刘奎龄	0.96	0.90	0.93	0.87	0.73	0.63
徐悲鸿	0.88	0.97	0.88	0.97	0.67	0.80
八大山人	0.97	0.93	0.93	0.93	0.67	0.60
平均值	0.94	0.94	0.91	0.91	0.68	0.67

实验进一步对实验结果及参数进行了讨论和分析,具体如下:

1) 实验中从每位画家的大量艺术作品中提取有效的艺术目标,并比较不同画家对于同一艺术目标风格描述的差异性以及稳定性,如图 8 所示。不难发现,徐悲鸿笔下的马盛气凌人,每一笔都渲染出马姿态的洒脱和傲立之感。而刘奎龄笔下的马在于写实,细致到马的每一缕鬃毛,颜色精致,仿佛照片中的马;程十发笔下的鸡往往成对出现,其形态憨态可掬,惟妙惟肖。而陈大羽笔下的鸡给人以傲气凌人,狂放不羁的感觉。发现艺术目标从人眼视觉角度考量具有内部一致性和外部差异性。由图 8 可以看出,徐悲鸿和程十发虽然在艺术目标的内容上与其他画家有所重叠,但其所表现出的艺术风格却大相径庭。在表 1 本文算法的案例 2 中,徐悲鸿和程十发都

获得了高达 97% 的查全率,也印证了以上的分析。

2) 如表 1 的案例 2 所示,本文算法中黄胄的查准率和查全率分别为 79% 和 73%,是十位画家中分类准确率最低的。这是由于黄胄在文革时期长期在北京故宫、琉璃厂临摹画作,博采众家之长,使得对于同一艺术目标其风格也变幻莫测。以其笔下的鸡为例,由图 9 可以看出黄胄所画的 3 只鸡分别运用了西方素描与中国笔墨相结合、泼墨、工笔的画法,其艺术风格的多变提高了计算机分辨黄胄画作中艺术目标的难度,降低了分类准确率。

3) 实验中还对不同艺术家相同题材的画作进行了研究,以山水画为例,如图 10 所示。唐寅笔下的山水画大多表现雄伟险峻的重山复岭,楼阁溪桥的江山胜景和文人逸士悠闲的生活。其笔下的“山石”皴法丰富,多有短砍、长皴、逆毫等笔墨,劲健峭



图8 4位画家相同艺术目标所展现出的不同艺术风格
Fig.8 Empirical study upon artistic styles of horses and chickens comparison across different artists ((a) Xu Beihong; (b) Liu Guiling; (c) Cheng Shifa; (d) Chen Dayu)



Fig.9 Chicken of Huang Zhou

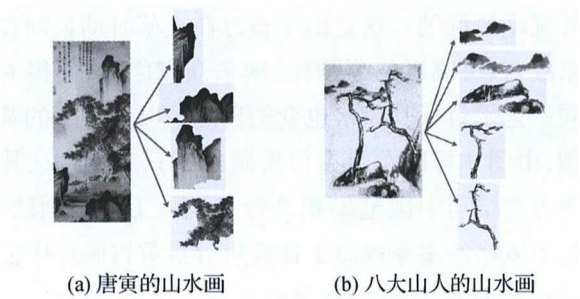


图10 不同艺术家的山水画风格迥异
Fig.10 Illustration of typical “Shan-Shui” painting
(a) Tang Yin; (b) Bada Shanren)

利,坚凝厚实。其笔下的“树”多为劲松,高大虬曲、清劲秀雅。而八大山人笔下的山水画多为荒山野

岭、残山剩水之景,给人一种荒凉残败、野逸无人之感,以表达其内心的悲愤与孤独。其笔下的“山石”多用渴笔手法,干笔涂抹,以求苍茫之韵。其笔下的“树”多为浓墨重点法,以求画面干湿平衡,树枝作下垂势,树干瘦硬如铁,树顶光秃。实验中发现,虽然唐寅和八大山人都擅长画“山水画”,但其所表现出的艺术风格却迥然不同。虽然其山水画中都描绘了山石、树等艺术目标,但每种艺术目标所展现出的状态和美感却截然不同。这一特点也在实验中得到印证,实验结果表明两位画家的山水画都能够得到正确的分类结果。

实验选取的10位画家的绘画题材涵盖了山水、花鸟、人物等领域,本文算法均能得到较高的分类准确率。对于国画中的某些元素,如流水等,由于其在中国画的表现中往往没有固定的轮廓、甚至只是采用留白的方式进行表现。例如,画水的一种典型画法是“留白法”,即有水的地方完全留白不加笔墨,以表现出水的存在,并且这种“诗情画意”的“留白法”被广泛的使用在山水画中。本文的分割算法很难提取出“水”等没有明显轮廓或过于抽象的艺术目标,在实验中也并没有考虑这种艺术目标。但是这并不影响本文对山水画进行分析和分类研究。通过提取画作中的其他关键艺术目标,如山水画中的山石、树、人物、房子等,已经足以判断画作和画家的艺术风格特点。

4) 实验进一步讨论了不同参数对结果的影响。图11表明,式(1)中 L 越小,分割后的超像素个数

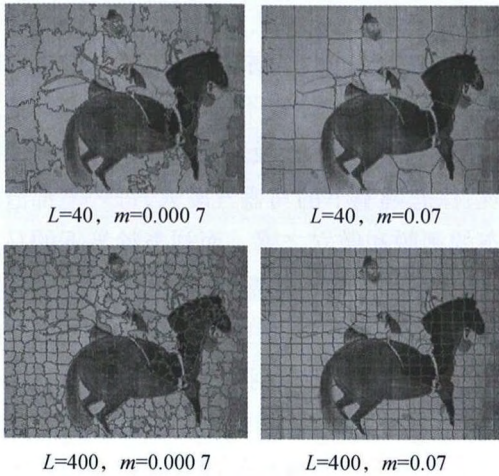


图11 不同参数对分割的影响
Fig.11 Performance comparison of object segmentation using different parameters

越少; L 越大, 分割后的超像素个数越多。式(4)中 m 越大, 空间位置距离越重要, 产生的超像素也更规整。 m 越小, 产生的超像素会更符合图像中物体的边界, 但是它的大小和形状却不规整。本文实验中, $L=400$, $m=0.07$ 。实验表明, 参数 m 和 L 的取值会在一定程度上影响图像分割的结果, 但对于画作整体的最终分类结果影响不大。

4 结 论

艺术目标如鸟、人物等, 是画家表达其思想感情的基本单元, 由于中国画艺术目标的有限性, 导致不同画家赋予相同艺术目标以不同的情感表达, 存在很大的研究空间。另一方面, 现有文献大多对整幅图像直接进行特征提取, 但任何图像内容特征的适应性都存在一定的局限性。如果不考虑每个笔锋所产生的条件而一味地分析笔锋的走向和力度分布等将会是很盲目的。因此, 本文提出一种最大相似度区域融合的中国画艺术目标提取算法, 将中国画交互式分割成一系列的艺术目标, 以提取画家用来表现艺术形式和抒发情感的相对稳定单元。为了更好的提取艺术目标的数字艺术风格特征, 本文通过微调 VGG-F 模型来提取特征, 该方法结合了 CNN 在特征提取方面的强大优势, 隐式地从训练数据中对这些艺术目标进行学习, 进而实现国画艺术目标的风格描述。最后, 引入支持向量机对每幅中国画内的各种艺术目标的分类结果进行最后的融合, 从而更好的实现中国画的分析与识别。

本文的研究成果可用于中国画的数字化分析、管理、理解和识别, 为中国画传承和鉴赏提供有效的数字工具。研究成果可转化应用到文化保护、建立数字图书馆、国画鉴赏、艺术培训等方面, 有利于传统文化的传承创新, 推动文化大发展大繁荣。

参考文献(References)

- [1] Qian W H, Xu D, Guan Z, et al. Simulating chalk art style painting[J]. Journal of Image and Graphics, 2017, 22(5): 622-630. [钱文华, 徐丹, 官铮, 等. 粉笔画艺术风格模拟[J]. 中国图象图形学报, 2017, 22(5): 622-630.] [DOI: 10.1142/S0218001417590261]
- [2] Sheng J C, Jiang J M. Recognition of Chinese artists via windowed and entropy balanced fusion in classification of their authored ink and wash paintings (IWP) [J]. Pattern Recognition, 2014, 47(2): 612-622. [DOI: 10.1016/j.patcog.2013.08.017]
- [3] Sheng J C. Automatic categorization of traditional Chinese paintings based on wavelet transform [J]. 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[J]. 计算机科学, 2014, 41(2): 317-319.] [DOI: 10.3969/j.issn.1002-437X.2014.02.069]
- [4] Li J, Wang J Z. Studying digital imagery of ancient paintings by mixtures of stochastic models [J]. IEEE Transactions on Image Processing, 2004, 13(3): 340-353. [DOI: 10.1109/TIP.2003.821349]
- [5] Jiang S Q, Huang Q M, Ye Q X, et al. An effective method to detect and categorize digitized traditional Chinese paintings [J]. Pattern Recognition Letters, 2006, 27(7): 734-746. [DOI: 10.1016/j.patrec.2005.10.017]
- [6] Sun M J, Zhang D, Wang Z, et al. Monte Carlo convex hull model for classification of traditional Chinese paintings [J]. Neurocomputing, 2016, 171: 788-797. [DOI: 10.1016/j.neucom.2015.08.013]
- [7] Wang Z, Sun M J, Han Y H, et al. Supervised heterogeneous sparse feature selection for Chinese paintings classification [J]. Journal of Computer-Aided Design & Computer Graphics, 2013, 25(12): 1848-1855. [王征, 孙美君, 韩亚洪, 等. 监督式异构稀疏特征选择的国画分类和预测[J]. 计算机辅助设计与图形学学报, 2013, 25(12): 1848-1855.]
- [8] Guo F, Peng H, Tang J. A novel method of converting photograph into Chinese ink painting [J]. IEEE Transactions on Electrical and Electronic Engineering, 2015, 10(3): 320-329. [DOI: 10.1002/tee.22088]
- [9] Liu J A, The Palace Museum. Image Catalog of Authenticity Identification for Chinese Paintings and Calligraphy [M]. Beijing: Forbidden City Press, 2013: 1-360. [刘九庵, 故宫博物院. 中国历代书画真伪对照图录[M]. 北京: 故宫出版社, 2013: 1-360.]
- [10] Guo X G. Judging Arts by Xiaoguang [M]. Beijing: China Renmin University Press, 2014: 10-77. [郭晓光. 晓光鉴画[M]. 北京: 中国人民大学出版社, 2014: 10-77.]
- [11] Song X Y, Zhou L L, Li Z G, et al. Review on superpixel methods in image segmentation [J]. Journal of Image and Graphics, 2015, 20(5): 599-608. [宋熙煜, 周利莉, 李中国, 等. 图像分割中的超像素方法研究综述[J]. 中国图象图形学报, 2015, 20(5): 599-608.] [DOI: 10.11834/jig.20150502]
- [12] Moore A P, Prince S J D, Warrell J, et al. Superpixel lattices [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Anchorage, Alaska, USA: IEEE, 2008: 1-8. [DOI: 10.1109/CVPR.2008.4587471]
- [13] Felzenszwalb P F, Huttenlocher D P. Efficient graph-based image segmentation [J]. International Journal of Computer Vision,

- 2004 , 59 (2) : 167-181. [DOI: 10. 1023/B: VISI. 0000022288. 19776. 77]
- [14] Shi J B , Malik J. Normalized cuts and image segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence , 2000 , 22(8) : 888-905. [DOI: 10. 1109/34. 868688]
- [15] Li H Y , W S P. Parallelization of Mean Shift image segmentation algorithm [J]. Journal of Image and Graphics , 2013 , 18(12) : 1610-1619. [李宏益 , 吴素萍. Mean Shift 图像分割算法的并行化 [J]. 中国图象图形学报 , 2013 , 18(12) : 1610-1619.] [DOI: 10. 11834/jig. 20131209]
- [16] Levinstein A , Stere A , Kutulakos K N , et al. TurboPixels: fast superpixels using geometric flows [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence , 2009 , 31(12) : 2290-2297. [DOI: 10. 1109/TPAMI. 2009. 96]
- [17] Achanta R , Shaji A , Smith K , et al. SLIC superpixels compared to state-of-the-art superpixel methods [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence , 2012 , 34 (11) : 2274-2282. [DOI: 10. 1109/TPAMI. 2012. 120]
- [18] Yang D X. Authenticity Identification of Chinese Paintings and Calligraphy [M]. 3rd ed. Shenyang: Liaoning People's Publishing House , 2016: 149-196. [杨丹霞. 中国书画真伪识别 [M]. 3 版. 沈阳: 辽宁人民出版社 , 2016: 149-196.]
- [19] Meyer F , Beucher S. Morphological segmentation [J]. Journal of Visual Communication and Image Representation , 1990 , 1(1) : 21-46. [DOI: 10. 1016/1047-3203(90) 90014-M]
- [20] Schmidhuber J. Deep learning in neural networks: an overview [J]. Neural Networks , 2015 , 61: 85-117. [DOI: 10. 1016/j. neutet. 2014. 09. 003]
- [21] Zhang X P , Xiong H K , Zhou W G , et al. Picking deep filter responses for fine-grained image recognition [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas , NV , USA: IEEE , 2016: 1134-1142. [DOI: 10. 1109/CVPR. 2016. 128]
- [22] Liu J , Gao C Q , Meng D Y , et al. Two-stream contextualized CNN for fine-grained image classification [C]//Proceedings of the 30th AAAI Conference on Artificial Intelligence. Phoenix , Arizona , USA: AAAI Press , 2016: 4232-4233.
- [23] He X T , Peng Y X. Fine-grained image classification via combining vision and language [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Honolulu , HI , USA: IEEE , 2017: 7332-7340. [DOI: 10. 1109/CVPR. 2017. 775]
- [24] Feng Y S , Wang Z L. Fine-grained image categorization with segmentation based on top-down attention map [J]. Journal of Image and Graphics , 2016 , 21(9) : 1147-1154. [冯语姗 , 王子磊. 自上而下注意力图分割的细粒度图像分类 [J]. 中国图象图形学报 , 2016 , 21(9) : 1147-1154.] [DOI: 10. 11834/jig. 20160904]
- [25] Chatfield K , Simonyan K , Vedaldi A , et al. Return of the devil in the details: delving deep into convolutional nets [C]//Proceedings of the British Machine Vision Conference. London , UK: BMVA Press , 2014: 1-12. [DOI: 10. 5244/C. 28. 6]

作者简介



盛家川 , 1982 年生 , 女 , 副教授 2013 年于天津大学获计算机应用技术专业博士学位 , 主要从事图像处理、模式识别方面的研究。
E-mail: jiachuansheng@tjufe. edu. cn

李玉芝 , 女 , 讲师 , 主要从事机器学习、多媒体处理方面的研究。 E-mail: liyuzhi@tjufe. edu. cn