面向中国古画鉴定的多模态可视分析方法

陈晓皎[1,2], 陈泳豪[2], 唐谈[1,2]*, 葛鑫[2], 汪睿涵[2], 王一帆[1], 王小松[1,2]

[1] (浙江大学艺术与考古图像数据实验室 杭州 310028)
[2] (浙江大学计算机科学与技术学院 杭州 310027)
(tangtan@zju.edu.cn)

摘 要: 中国古画的传统鉴定很大程度上依赖于专家的推理论证, 不仅需要考虑题款、印章、画作内容等多方面的信息, 还需大量的文献资料辅助. 基于深度学习的方法提高了传统鉴定工作的效率, 但其主要聚焦于图像分析, 同时深度神经网络的黑盒特性不利于人机高效协作. 为此, 提出了一种多模态语义提取的可视分析方法. 首先将古画图像分割为题款、印章、画作内容 3 类语义对象; 接着通过图像匹配和文本识别构建对应古画之间及与相关文献条目的关联结构并编码; 最后由专家进行古画片段的选取以生成多模态语义网络关系图进行查阅考证; 此外, 设计并实现人机协同驱动的中国古画辅助可视鉴定系统. 通过案例研究和专家评估并以 5 点李克特量表量化专家反馈, 得分范围在 0.78~1.78, 验证了所提方法在古画分析和考证关联中的高效性和准确性.

关键词: 中国古画; 鉴定; 计算机辅助; 多模态; 可视分析
中图法分类号: TP391.41 **DOI:** 10.3724/SP.J.1089.2023-00350

Multi-Modal Visual Analytic Approach for Attributing and Authenticating Ancient Chinese Paintings

Chen Xiaojiao[1,2], Chen Yonghao[2], Tang Tan[1,2]*, Ge Xin[2], Wang Ruihan[2], Wang Yifan[1], and Wang Xiaosong[1,2]

[1] (*Laboratory of Art and Archaeology Image, Zhejiang University, Hangzhou* 310028)
[2] (*College of Computer Science and Technology, Zhejiang University, Hangzhou* 310027)

Abstract: The traditional workflow of authenticating ancient Chinese paintings heavily relies on experts' reasoning and justification, who need to consider various information, such as inscriptions, seals, and painting content, along with extensive literature materials. Although deep learning-based methods have improved the efficiency of traditional authentication workflow, they predominantly focus on image analysis while the black-box nature of DNNs hampering the effective human-machine collaboration. To address this issue, we propose a visual analysis approach that extract multi-modal semantics to facilitate the easy analysis of ancient Chinese paintings. Firstly, ancient painting images are segmented into three semantic objects: inscriptions, seals, and painting content. Then, an association structure is constructed between the corresponding segments of ancient paintings through image matching and text recognition, as well as with relevant literary references. Finally, experts select segments of ancient paintings to generate a multi-modal semantic network graph for reference and verification. Furthermore, we develop a mixed-initiative system for assisting the visual attribution analysis of ancient Chinese paintings. Through case studies and expert evaluations, the effectiveness and accu-

收稿日期: 2023-06-18; 修回日期: 2023-10-10. **基金项目**: 国家社会科学基金重大项目(19ZDA046); 浙江省哲学社会科学重点基金(22JCXK03Z). **陈晓皎**(1988—), 女, 博士, 研究员, 博士生导师, 主要研究方向为数字人文、人机交互; **陈泳豪**(2000—), 男, 博士研究生, 主要研究方向为智能交互; **唐谈**(1994—), 男, 博士, 研究员, 博士生导师, 论文通信作者, 主要研究方向为数据可视化; **葛鑫**(1994—), 男, 博士研究生, 主要研究方向为可视化; **汪睿涵**(2001—), 男, 硕士研究生, 主要研究方向为可视化; **王一帆**(1984—), 男, 博士, 研究员, 主要研究方向为清代绘画鉴别; **王小松**(1964—), 男, 博士, 教授, 博士生导师, 主要研究方向为艺术与设计.

racy of ancient painting analysis and authentication were validated using a 5-point Likert scale for expert feedback, with scores ranging from 0.78 to 1.78.

Key words: ancient Chinese painting; authentication; computer-aided; multi-modal; visual analytics

中国古代绘画(以下简称中国古画)的真伪鉴定在中外艺术史上一直是一个很重要的研究领域. 随着信息技术的发展, 数字化的中国古画图像已成为主流展示形式, 许多研究工作开始围绕数字化图像展开[1-2]; 计算机辅助鉴定是当下的研究热点[3-4], 但是因为算法黑盒特性等因素导致其实际应用转化存在一定的局限[5].

中国古画的传统鉴伪工作需要大量的领域知识和经验, 并同时考虑题款、印鉴和技法等要素[6], 故局限于单一对象分析的计算机技术是无法满足需求的. 计算机技术介入古画鉴定分析时同样需要考虑鉴定涉及的多类对象, 这对计算机辅助鉴定研究提出更高的要求. 除了对古画本身进行分析, 外延的文献资料考证也是必不可少的工作, 但有效的信息调研过程往往耗时、耗力[7]. 此外, 目前以机器学习和深度学习方法为代表的书画鉴定工作普遍是在特定数据集下进行训练并直接给出判定结果[8-10], 其中关键的图像深层特征具有不可解释性, 也无法和专家进行交互与验证. 因此, 在古画鉴定工作中, 计算机技术应该定位于为专家提供辅助, 如捕捉画面内容的相似片段并交予专家判断[11-12].

为应对上述挑战, 本文设计了中国古画多模态语义分析鉴定框架.

(1) 运用多模型联合技术对画作图像进行处理, 使用通过自监督方法预训练完成的万物分割模型(segment anything model, SAM)在中国古画图像数据集中实现零样本无监督的迁移泛化, 以达到对题款、印鉴等目标的精准语义分割, 再为题款、印鉴和内容 3 类语义片段训练不同的检测模型, 实现信息的精准提取和内容的高效匹配.

(2) 基于收集整理的古画文献资料库, 提取出关键考证信息并与相应的考证对象关联, 预先构建关系结构数据, 定义了跨文本、图像、关系结构 3 类数据的中国古画鉴定多模态语义, 由此外延的考证信息可以依托关系网络结构快速检索.

(3) 设计了一个中国古画计算机辅助鉴定系统, 允许专家以交互方式自由调用深度学习模型对画作图像进行分割切片和匹配. 系统能够自动量化画作切片的关联信息并进行可视化, 同时构建关联信息的关系网络; 专家可以通过关系网络查阅与选中片段具有信息关联的多模态语义数据, 以此为专家的鉴定工作提供辅助支持.

最后, 本文进行了一项中国古画鉴定专家参与的案例研究及专家评估, 以分析本系统在中国古画鉴定过程中辅助专家的实际效益.

1 相关工作

1.1 传统鉴定方法研究

相较于西方绘画艺术, 中国古画更为抽象. 例如, 西方绘画鉴定中使用的颜色直方图特征在中国古画中并不适用[13], 这为鉴定中国古画增加了难度. 鉴定古画的证据通常来自 3 个方面, 包括文字材料、图像资料和物证材料. 本文研究不局限于图像的特征提取, 还纳入了文献等考证材料.

在中国古画史中, 石涛是著名的艺术家之一, 他的画作价值极高. 由于石涛生活在一个造伪盛行的时代, 因此, 临摹石涛的绘画伪作最多, 已有许多学者专注于石涛绘画的真伪鉴定研究. 这些研究虽然已经取得一定成果, 但仍有许多绘画鉴定存在分歧, 尤其是张大千对石涛的模仿极为相似, 真伪难辨[14]. 石涛的绘画在艺术上具有独创性, 具体体现在其绘画技法中笔法点与圈的应用[15]. 古画鉴定的另一项挑战在于能够比对考证的图像资料极少, 这也同样导致计算绘画领域中十分重视绘画数据集的整理[16]. 历史上大部分中国画画家留存的作品数据较少, 本文收集了中国历朝历代共 12 405 幅传世绘画作品作为中国古画图像数据集, 其中依旧不乏难以辨别的相似画作和一位作者仅有一幅作品流传的情况. 所有画家存世的作品数量中以石涛为最, 粗略估计超过 500 幅, 最适合总结归纳普遍规律特征. 因此本文选择了石涛的绘画作品并尝试借助计算机技术进行辅助鉴定.

1.2 数字化图像检索系统研究

近年来, 绘画的数字存储为绘画比较、搜索和研究带来便利. 同时, 数字图像数据有利于采用计算机技术进行可视分析. 为了提升图像处理的效率, 降低人工成本, 使用计算机技术分析是必然的选择, 包括物体识别、相似性计算和分类等. 20 世纪 70 年代末已经出现基于文本描述的图像检索技

术, 90 年代后开始转向开发基于颜色、纹理等内容特征的分类和检索系统[17-18]. 例如, Pflüger 等[19]开发了一个大型图像集的可视化系统, 它能快速地计算距离函数、聚类和投影. 除了图像检索系统, Gleicher[20]开发了能比较绘画细节的可视化系统, 包括并排比较(并列)、混合(叠加)和明确的差异编码(聚合), 或是为优化可视化系统而融合多种比较方式. 例如, Pflüger 等[12]介绍了一种探索性的视觉方法, 他们提供放缩和切换等功能, 以促进快速且准确地支持比较版画之间的细微偏差.

图像系统的可视分析还应当基于多种要素考虑, 如画面的线条提取[21]、像素的聚类分析[22]和画作的风格感知[23]等. 然而每种方法都有它的局限性, 且针对不同的分析对象使用的方法也不尽相同. 基于此, 设计多模态的可视化分析方法从多个角度进行完善的分析是有必要的.

1.3　计算机技术鉴别研究

为了弥补人眼鉴别绘画时的局限, 一些学者结合包括计算机技术、光电技术和化学成分分析等技术, 以提供辅助鉴别的信息. 光谱成像和光谱分析被认为在绘画鉴别方面有独特优势[24]. Xu 等[25]认为, 光谱成像和分析对于颜料的识别、隐藏信息的提取、弱信息的增强以及不同性质的分类和识别都很有用, 物理化学方法也常用于对画作年代的检测[26].

然而, 随着研究深入, 这些技术已经不能满足更细致的鉴别需求. 机器学习的发展带来了更多特征的检测方法, Swain 等[27]最早提出颜色直方图概念, 对颜色特征进行提取分析, Stricker 等[28]和 Hsu 等[29]的研究深化了颜色特征; Haralick 等[30]通过灰度关系构造了一系列统计特征作为纹理特征, Tamura 等[31]将纹理特征映射到人的视觉感知集中, 尺寸不变特征变换[32]和小波变换[33]等方法也是当前图像常用的特征提取方法; 更进一步地, 基于形状轮廓特征[34]、笔墨特征[35]、群体智能[36]等方法都能在画作识别分类上起到一定的作用.

深度学习的兴起, 使得研究人员转而使用卷积神经网络(convolutional neural network, CNN)来提取画作的特征. 与传统方法相比, 深度学习方法能够提取到更深层次的特征, 且特征表达更准确. 例如, 李大湘等[8]使用多尺度 CNN 提取国画特征, 在 CNN 的基础上, 文献[9-10]加入信息论和目标分割, 使得分类更精准. 同样在西方绘画的研究中, Shen 等[11]基于空间一致的特征学习方法发现艺术品数据集中的重复区域; Elgammal 等[37]使用深度学习提取线描画笔画特征.

2　背景与需求分析

从机器学习到复杂的深度学习模型, 算法模型有了一系列的改进, 但是这些改进在画作鉴伪应用中的适用情况和实质性的推进作用仍然是非常有限的[5]. 本文对 9 位专业的中国古画鉴定专家组成的专家小组进行了针对性的采访, 目的是确定他们鉴定工作中的核心需求, 以找出一个系统必须提供哪些功能和属性来支持这项工作. 专家访谈表明, 中国古画的鉴定过程需要同时考虑画作中的题款、印鉴以及绘画内容中蕴含的笔触技法色彩等要素, 目前大部分研究往往基于画作内容的某一方面作为切入点, 这是不全面的. 随着技术的改进, 提取的特征也越来越复杂, 开始变得不可解释, 使专家难以理解和分析. 此外, 模型的设计都倾向于直接给出鉴定的推断结果, 造成了一定的人机矛盾. 最后, 专家提到相关研究基本上都是围绕分析画作图像本身开展的, 但是鉴定的过程中需要大量的文献资料辅证, 当前大部分情况下鉴定都以考证为主, 最耗费精力的工作则是资料的查找.

经过对鉴定专家小组的访谈, 本文梳理了专家对当下画作鉴定系统的几个要求.

R1. 同时针对多个对象处理. 制伪者已经掌握了高超的伪造技法, 并且尽可能地模仿真品的构图、笔法、墨色和色彩等特征. 一些制作者可能会伪造和使用真品的题款或印章, 或者使用历史上存在过的画家名字, 这些假题会让鉴定专家更难以判断真伪. 还有一些制作者可能会将多幅不同真品画作的不同部分组合成一幅伪作, 由此鉴定分析的对象一定要足够全面, 覆盖整幅画尽可能多的要素.

R2. 选择适当的模型和可视设计方法. 系统中使用的模型应当产生直观易懂的结果, 而不是生成过于复杂晦涩的特征图样, 同样也不能选择直接输出预测的模型, 因为这将对专家深度协同造成阻碍. 因此, 系统需要设计人机友好的可视化界面, 并且基于专家需求支持二次交互.

R3. 能够快速地检索与关联匹配关联信息. 鉴定过程中的文献资料查找比对是最重要的步骤, 大部分计算机画作分类或鉴定工作都忽视了考证数据的重要性, 而信息快速检索正是计算机相比人力的优势. 系统除了应当及时地将提取到的文本、图像等信息与图像数据库中的资料进行比较并反馈给专家, 更应当能够快速地关联相关的外延

文献考证资料进行展示.

基于计算机技术的辅助真伪鉴定目前有 2 种模式[3]. 一是对识别对象自动真伪鉴定或是给出一些量化指标, 近年来大量的鉴定研究都是围绕特定画家的绘画数据集训练一个特化的自动鉴定模型, 但这样的方法不具有泛化性. 二是对要鉴定的对象进行辅助性分类和样本搜索, 这需要大量数据集作为支撑. 因此新的辅助鉴定系统设计应当兼顾图像自动处理需求和相关内容的匹配搜索, 还应设计友好的人机协同用户交互界面.

考虑鉴定涉及的要素过多, 首先对需要处理的对象进行统一的整理和定义(R1). 在形式上, 题款、印鉴和画作内容是最具有代表性的 3 类对象. 题款是画作中的文本数据; 印鉴和画作内容同属于图像数据, 但是两者还是有明显的不同. 印鉴来自一个固定印章, 其在不同的画作上大同小异; 画作内容是作者在画布上绘制的不同物象, 蕴含的是作者的笔墨技法色彩等特征. 另外, 本文还重点考虑了文献考证信息的需求纳入(R3), 上述 3 类对象在史料中可能存在的关联辅证, 会在很大程度上影响最终的鉴定结果. 3 类对象彼此与文献资料之间的考证关联构建了关系结构数据, 故而完善的中国古画鉴定分析框架中的数据形式应是多模态的. 本文定义多模态语义鉴定框架中古画数据集 $P=\{p_1,p_2,\cdots,p_n\}$ 包含 n 幅作品, 第 i 幅作品 p_i 包含题款 $T_i=\{t_1,t_2,\cdots,t_{x_i}\}$, 印鉴 $S_i=\{s_1,s_2,\cdots,s_{y_i}\}$ 和画作内容 $C_i=\{c_1,c_2,\cdots,c_{z_i}\}$ 这 3 类对象. 其中 x_i 为第 i 幅作品的题款数目, y_i 为第 i 幅作品的印鉴数目, z_i 为第 i 幅作品中除题款和印鉴外的内容要素数目. 此外 $H=\{h_1,h_2,\cdots,h_m\}$ 代表从历史文献中整理的考证资料条目, 它们记录了与古画、题款、印鉴、内容技法相关的信息, 如印鉴出现在文献中的使用记录、画家不同时期的作画技法特色记录、画作的流传收藏记录等. 要素之间关联可表示为 $(P,T,S,C)\rightarrow H$, 即上述要素 P,T,S,C 中与 H 之间的关联; 进一步所有对象要素均可视为节点 V, 与其余要素之间的信息关联抽象为连边 E, 构建关系网络 $G=\langle V,E\rangle$, 并在后续的分析过程中继续扩充此关系网络.

本文的中国古画多模态语义分析鉴定框架如图 1 所示.

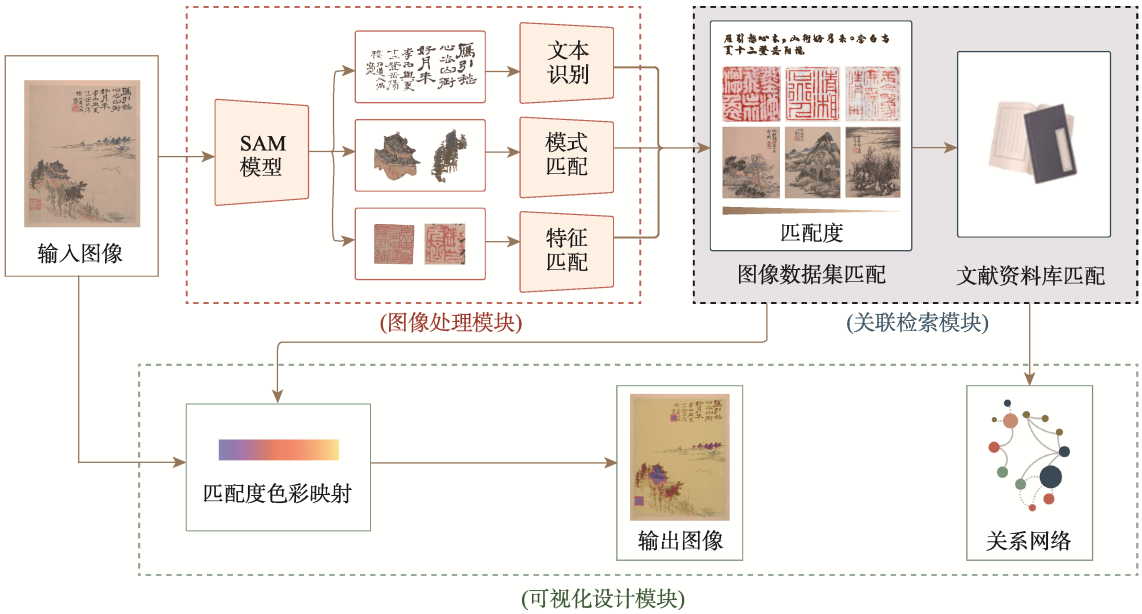


图 1 多模态语义分析鉴定框架

3 数据处理模型

本文系统首先选择合适的模型搭建多模型联合模块对输入的图像进行处理(R2), 包括 1 个图像语义分割模型和 3 个针对不同语义对象后续处理的特定模型.

在语义分割的工作上, 本文基于 Meta 最新的 SAM[38]实现, SAM 在 1 100 万幅公开或私有的图像上经过预训练, 可以零样本直接泛化到新的图像任务中, 在中国古画的语义分割上也取得了图 2 所示的优良效果, 分割片段可直接作为后续模型的输入. 在交互方面, 本系统通过微调输入的参数格

式，整合了点击识别、矩形框选识别和全图识别 3 种交互模式，识别的图像区域可以自由选择(R2). 具体地，基于分割的任务需求，SAM 支持的提示符在编码方式上也有所不同. 点、矩形(左上点和右下点)采用极坐标系下的 sin 和 cos 值与特征点的类别进行编码，而全图采用预训练的 Vision Transformer 和 U-Net 方式进行编码.



a. 分割前　　b. 分割后

图 2　《竹林致深》SAM 测试

题款的鉴定常被抽象为签名的验证方法，主要包含静态特征[39-40]和下笔过程中的动态特征[41]. 而实际场景中，题款笔迹受笔墨和印鉴重叠的影响，并不像书法作品一样适合通过单字字迹来鉴定真伪，专家更多是通过解读文本及文本内容的语义信息来考证. 因此题款的处理可以选择文字识别模型，通过尽可能地识别出一些有意义的关键词组，并将其作为输入在考证资料中检索出有价值的资料，以产生更加实际的效用.

在文本片段的识别上，本文采用了在光学字符识别(optical character recognition, OCR)中表现较好的 TransOCR 模型[42]. 它是一种基于 Transformer 的文本识别方法[43]，由一个基于 CNN 的文本超分辨率框架场景文本望远镜(scene text telescope, STT)和一个基于 Transformer 的序列解码器组成. TransOCR 模型能够有效地捕捉文本图像中的上下文信息和长距离依赖关系，从而提高识别准确率. 因为只有部分古画中有题款的数据样本，样本还可能和画中的笔墨、印鉴部分重叠而影响样本质量，题款字体风格也因落款人和时代不同有很大差异. 为了使 TransOCR 能够更好地适应中国传统绘画的题跋部分的文本识别任务，本文使用 Liang 等[44]发布的一个关于中国繁体手写的数据集进行训练样本扩充. 该数据集是目前公开的第 1 个专门针对中国繁体手写书法的数据集，包含来自不同书法家、不同年代、不同风格的中国繁体手写书法图像，共有 672 163 幅图像和 5 301 个文字标签，使得 TransOCR 模型能够更好地学习题跋部分的特征和规律，从而提高识别准确率和鲁棒性. 识别所得到的文本作为字符串检索算法的输入在考证资料库完成关键文本信息匹配. 通过该部分工作可以在网络 G 中扩充任意题款和画作及考证条目之间的关联 $T_i \to p_i$ 和 $T_i \to H$.

在印鉴鉴定方面，实际场景中由于加盖印鉴的条件不同，本就会造成同一印章的不同印鉴之间存在较大差异，这种差异的存在使真假印鉴的区分变得十分困难[45]；模型只能够基于相似性对印鉴进行初步匹配，然后需要更多外延资料的考证. 在印鉴匹配模型的训练上，本文将所有的印鉴分割片段进行整理作为印鉴数据集，将 VGG-16 网络和 Siamese 网络相结合，使用 Contrastive Loss 作为损失函数，训练任意 2 个印鉴图像的相似度判断模型，提取 16 维特征向量代表输入印鉴图像的特征. 假设 2 个输出的 16 维印鉴图像特征向量分别为 $\boldsymbol{v}_i$ 和 $\boldsymbol{v}_j$，则它们的相似度可以表示为

$$\cos\theta = \frac{\boldsymbol{v}_i \cdot \boldsymbol{v}_j}{\|\boldsymbol{v}_i\| \cdot \|\boldsymbol{v}_j\|} \tag{1}$$

由此可以在网络 G 中扩充印鉴与画作之间的来源关系 $S_i \to p_i$ 和古画数据集中相似印鉴的关联 $S_i \to S_j$.

在绘画内容上，每个作者都有各自独特的风格，包括绘制的物象[46]、色彩的使用[47]和不同阶段的风格变化[48]等，这要求模型能够自动地捕获图像数据集中不同图像间相似的区域. Shen 等[11]和 Ufer 等[49-50]在图像数据集中寻找相似区域对象上已经有了成功的实践，以完整图像作为处理单位，将提取的内容对象分割片段作为输入，可以大幅提高搜索效率. 本文结合区域投票算法和近似邻近搜索算法，将内容分割片段依据片段的类别分成 k 个子集，并使用区域投票算法找到每个子集中的代表向量. 对于待匹配的分割片段，使用近似邻近搜索算法来找到数据集中最相似的区域，再利用区域投票算法来确定该区域属于哪个子集，并进一步确定属于哪些图像. 第 i 个图像分割片段数据子集 D_i 的代表向量 $\boldsymbol{R}_i$ 表示为

$$\boldsymbol{R}_i = \arg\max_{\boldsymbol{x}\in D_i}\left(\min_{\boldsymbol{y}\in D_i,\boldsymbol{y}\neq\boldsymbol{x}} d(\boldsymbol{x},\boldsymbol{y})\right), \tag{2}$$

$$\arg\min_{i=1}^{k} d(\boldsymbol{q},\boldsymbol{R}_i) \tag{3}$$

其中，$d(\boldsymbol{x},\boldsymbol{y})$ 表示任意 2 个图像分割片段的特征向量 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 之间的欧几里得距离，并当 $\boldsymbol{q}$ 为待匹配片段的特征向量时从 k 个子集中确定最相似的子集以进一步比对. 以此在网络 G 中扩充内容分割片段与画作之间的来源关系 $C_i \to p_i$ 和古画数据集中相似内容片段的关联 $C_i \to C_j$.

4 系统设计

通过图像处理模块能够构建一定规模的鉴定语义对象关联结构，最终完善了以语义对象为节点，关联关系为连边的网络 G. 通过网络中节点和连边之间的跳转，实现了信息的快速关联(R3). 最后在系统视图设计上，需要合适的可视化方案来对一幅待鉴定作品进行展示和比对(R2,R3).

本系统界面基于 Vue 框架开发，整体运行效果如图 3 所示，主窗口部分可以导入任意待鉴定画作图像并实时展示，同时调用 SAM 进行预处理. 假定待鉴定画作为 U，经模型的处理后可提取得到题款 T_U、印鉴 S_U 和画作内容 C_U 这 3 类分割片段.

系统默认调用的是全图分割模式，可以在主视图下方的分割结果视图预览模型分割出的所有类型的语义对象. 但是考虑全图分割模式可能存在误差过大等问题，系统还调用了模型的点击模式和矩形框选模式接口，如果专家对于某一区域感兴趣，可以通过自由点击或者矩形框选的形式，使模型的分割部分能够尽量拟合目标区域，如图 3 主窗口所示.

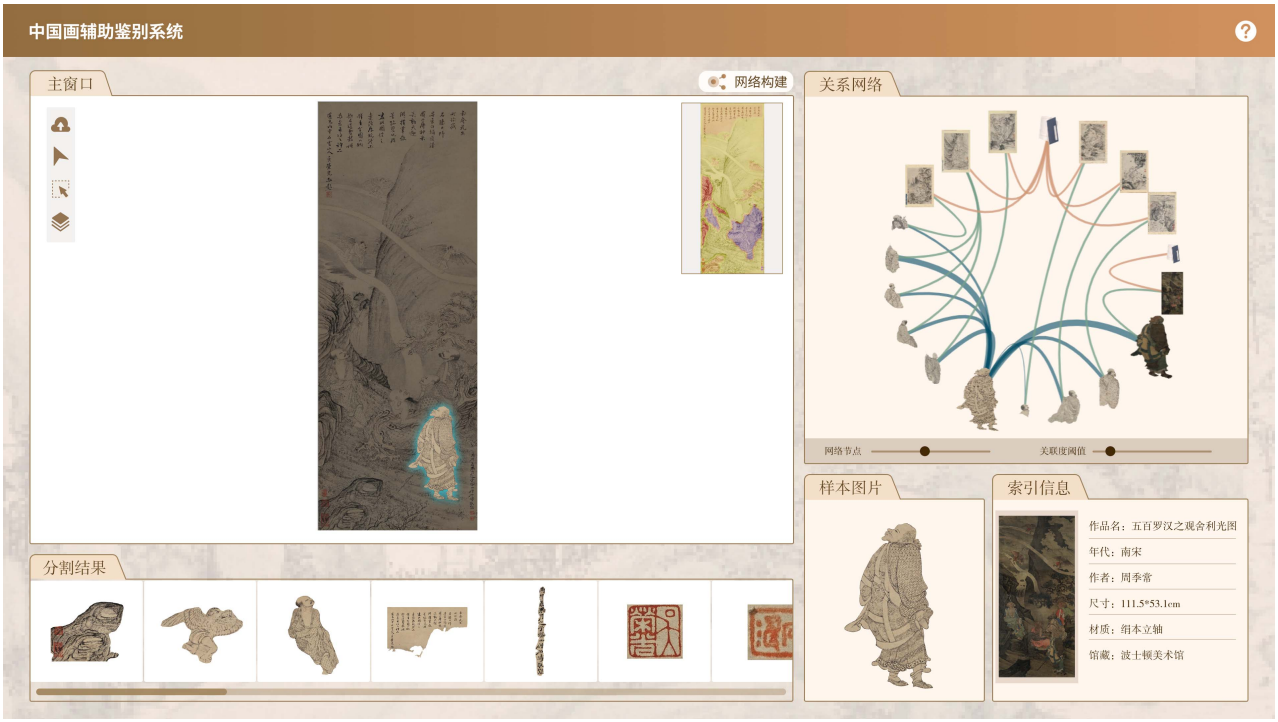
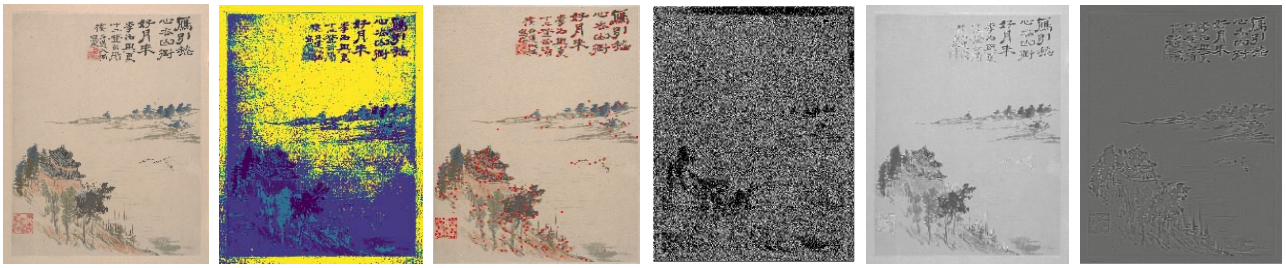


图 3 中国古画辅助鉴定系统界面

为了满足交互过程中的可理解性需求，本文使用了 2 种可视化方式. 首先是针对绘画作品本身，一直以来，图像特征的可视化解释都是研究的难点之一，尤其是在绘画鉴定领域. 从图 4 可以看到，从一般的提取特征到 CNN 的特征，都很难将可视化后的特征信息联系到专家在古画鉴定工作中所能够捕捉并解释的细节；甚至在大部分情况下，图像特征所表现的差异在原图像中也能直观地观察到. 因此，目前的特征都难以为绘画鉴定工作提供有效的解释和支持.

为此，本文所采用的可视设计之一是在绘画作品上叠加色彩映射. 不同的色彩区块直接反映对应分割的语义对象切片在数据库中所能够关联的信息条目数量，以此直观地向专家传达古画中值得分析关注的区域. 每类分割片段可以通过对应的模型与数据库的元素进行匹配，从而构建出覆盖题款、印鉴、内容、画作和考证信息的鉴定关系网络. 不同的分割片段能够关联的信息量是不



a. 原图　b. 色彩聚类[12]　c. 特征点检测[32]　d. 小波变换[33]　e. 边缘增强[30]　f. CNN[8]

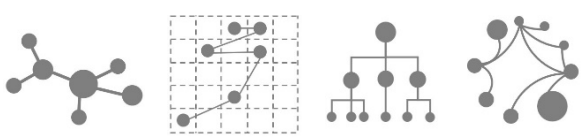
图 4　各类特征提取方法效果图

一样的，为了直观地可视化画作中切片的信息量，采用了不同的颜色直接映射的方式．切片信息量为关联节点数的累加，如画作 U 的题款 T_U 部分中某个子元素 t_i 关联节点数为 $\sum_{j=1}^{V_{\text{count}}} t_i \to v_j$ ．

V_{count} 为图 G 中所有节点的总数，v_j 为其中任意一个节点且属于 5 类语义节点 T,S,C,P,H 之一，表示 t_i 与 5 类语义节点 T,S,C,P,H 构建的关联连边总数，叠加了色彩映射的实际效果图展示在图 3 主窗口的小图中．

另一个可视设计针对每一个分割的语义对象，即不同色彩区块对应的详细关联网络，需要通过关系图的形式在关系网络视图中展现，并且关系图的可视化选择需要符合专家的理解和鉴定网络的结构特点．如图 5 所示 4 种布局，在网络原始结构的基础上，力导向布局通过节点之间的斥力和连边之间的引力进行迭代布局，因为没有对连边进行一定的约束，当数据量增长时最容易造成视觉上的连边交叉；另外，由于要反复迭代直至网络达到平衡状态，因此在计算资源上开销最大．网格布局对数据原始结构要求较高，它适用于规则结构化数据．树状图严格按照树形结构布局，易于理解和导航，但是本工作中节点之间可能生成闭环关系，与树状结构的实现逻辑产生冲突．故最终选择圆形布局，能够对节点的位置进行限制，使得每类节点沿圆弧规律排布，在较好聚类的同时，也尽量避免连边交叉．

在此基础上构建网络时，还应当同时考虑网络的语义节点数量规模和节点之间的关联度的影响，使得整体可视分析效果达到最佳．如图 6 所示，在节点数 V=11，连边数 E=12 的关系图中，当可视化数据量增加时，连边会和节点数量一并增多并出现交叉，不可避免地会带来视觉的混乱．同时节点之间关联度体现为连边的粗细，在固定节点数的前提下，可以通过过滤低关联的连边来二次优化视觉效果．在本文中，关联节点之间的关联度可以用语义处理模型的参数来定义，其中文本衍生的连边为有效地识别文本数和文本总数的比值，印章和内容节点则为模型推理的相似概率值．通过补充节点数量阈值和关联度阈值控制的过滤方式，既能够传达必要的关键信息，也能够减少视觉上的混乱．

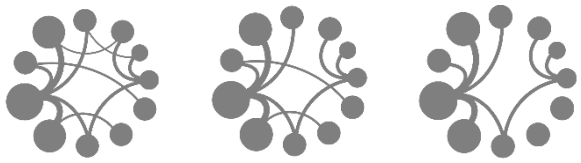


a. 力导向布局　b. 网格布局　c. 树形布局　d. 圆形布局

图 5　关系图布局类型



a. V=7　b. V=9　c. V=11



d. E=12　e. E=10　f. E=8

图 6　节点数量和连边优化对布局的影响

最后，为了满足专家使用过程中的对比需求(R3)，设计了图 3 右下方的对比视图．当专家选定需要构建关系网络的分割对象后，在对比视图的样本图像窗口显示该对象，同时可视分析视图生成对应的关系网络；专家可以任意查看该对象能够衍生到的节点，点击节点则节点信息将会在对比视图的索引信息窗口显示，以辅助专家进行详细的比对分析．

5 案例研究与系统评估

本文预先准备了未录入图像数据集的石涛画作作为测试案例, 邀请了 9 位鉴定领域的专家进行案例研究和系统评估. 通过案例研究评估本系统在鉴定工作中是否达到接近实际鉴定过程的预期效用; 通过专家体验对比本系统与传统鉴定方法的差异; 最后, 与专家围绕本系统和计算机辅助鉴定技术等一系列问题进行深度探讨.

本文收集了 484 幅石涛的画作, 并以图 7 所示《罗汉四条屏》第 1 屏为测试案例, 结合预处理的含 12 405 幅历代中国画数据集和考证资料库来测试本系统的实际效果.



图 7 《罗汉四条屏》

5.1 案例研究

在案例研究中, 其中 1 位专攻于石涛研究的专家作为代表被邀请进行测试. 专家首先基于自己的从业经历给出《罗汉四条屏》的关键考证条目, 并使用本系统来对《罗汉四条屏》第 1 屏进行鉴定分析, 最后将能够从系统中提取的有效信息与先验条目进行比对, 统计是否有相似条目. 其操作流程如下:

Step1. 导入待鉴定画作图像.

Step2. 等待系统自动进行全局语义分割和切片关联可视化处理. 待处理完毕, 系统会自动把全局的语义分割切片展示在下方陈列视图中, 同时基于不同切片所关联到的信息量叠加色彩映射, 将其添加到主窗口的小窗中.

Step3. 专家可以选择预生成的信息映射图来查看所能关联到的网络. 如果认为系统处理得不够严谨, 可以通过菜单提供的点击或者框选工具自主选定合适图像区域, 点击网络构建系统会针对用户自定义的选择再次处理.

Step4. 当专家浏览可视分析窗口时, 每个节点代表着关联的题款、印章、内容、画作和考证, 通过点击节点使得节点的详细信息在比对窗口显示, 比对窗口的左窗口是原始片段图样, 右窗口是节点信息. 通过比对或者新信息的补入, 对专家的考证过程起到推动作用.

Step5. 对于同一幅作品, 继续执行 Step3 和 Step4, 直到专家认定已经能够得到作品真伪的初步定论; 若要处理新的画作图像, 继续执行 Step1~Step4.

通过对图像内容的多次迭代关联考证, 专家能够从衍生的关联信息中推导并记录一些非常具有价值的信息, 实验结束后将记录条目与先验条目比对, 得到了一些有价值的重复条目并进行如图 8 所示的归纳.

5.2 专家评估

为了验证与评估中国古画计算机辅助鉴定系统相比于传统鉴定方式的有效性, 本文开展了一项面向中国古画鉴定的专家评估实验, 以石涛绘画鉴定为任务, 对比依据经验的鉴定方式与计算机鉴定系统辅助的专家鉴定方式下的任务绩效与主观感受. 实验招募的 9 名(男性 7 名, 女性 2 名)拥有丰富鉴定经验的专家用户, 年龄在 45~63 岁, 视力或者矫正视力正常, 无色盲色弱症状. 实验使用尺寸为 23.8 英寸(约 60.45 cm)的并且分辨率为 2K(2 048×1 080)像素的显示器, 实验开始前进行校色处理. 正式实验开始前, 被试经过指导并练习操作实验系统, 确保被试充分了解中国古画可视辅助鉴定系统功能及交互过程, 并能够熟练使用.

在专家评估过程中, 专家被要求独立鉴定和运用计算机辅助鉴定系统进行绘画内容识别及真伪推断, 实验任务为识别画作题款、印章及绘画内容, 并填写主客观评估量表. 正式实验流程分为 3 个步骤.

Step1. 专家用户独立进行实验画作的内容识别, 写下题款文本、印章信息及绘画内容特征, 并通过自身经验联想相关辅证信息进行真伪推断.

Step2. 专家用户运用中国古画计算机辅助鉴定系统对当前实验画作进行评估, 考察辅助系统提取的题款、印章及绘画内容准确性, 以及系统提供的考证关联匹配信息准确性; 评估完成后填写准确性客观评估量表, 并比较专家用户联想信息的条目数均值和系统关联考证条目数, 将它们作为客观指标.

Step3. 专家用户对比 2 种鉴定方式(独立鉴定与计算机辅助鉴定)下识别任务的有效性, 包括对鉴定关键信息(题款、印章及绘画内容)的识别速度、识别准确性和考证关联信息高效性 3 类指标的量表打分.

实验采用 5 点李克特量表, 中位数为 0, 评分范围从−2(强烈不同意)到 2(强烈同意), 专家根据自己的主观感受进行填写. 问卷采用对比的问题设定方式, 如“您认为计算机辅助系统对题款的识

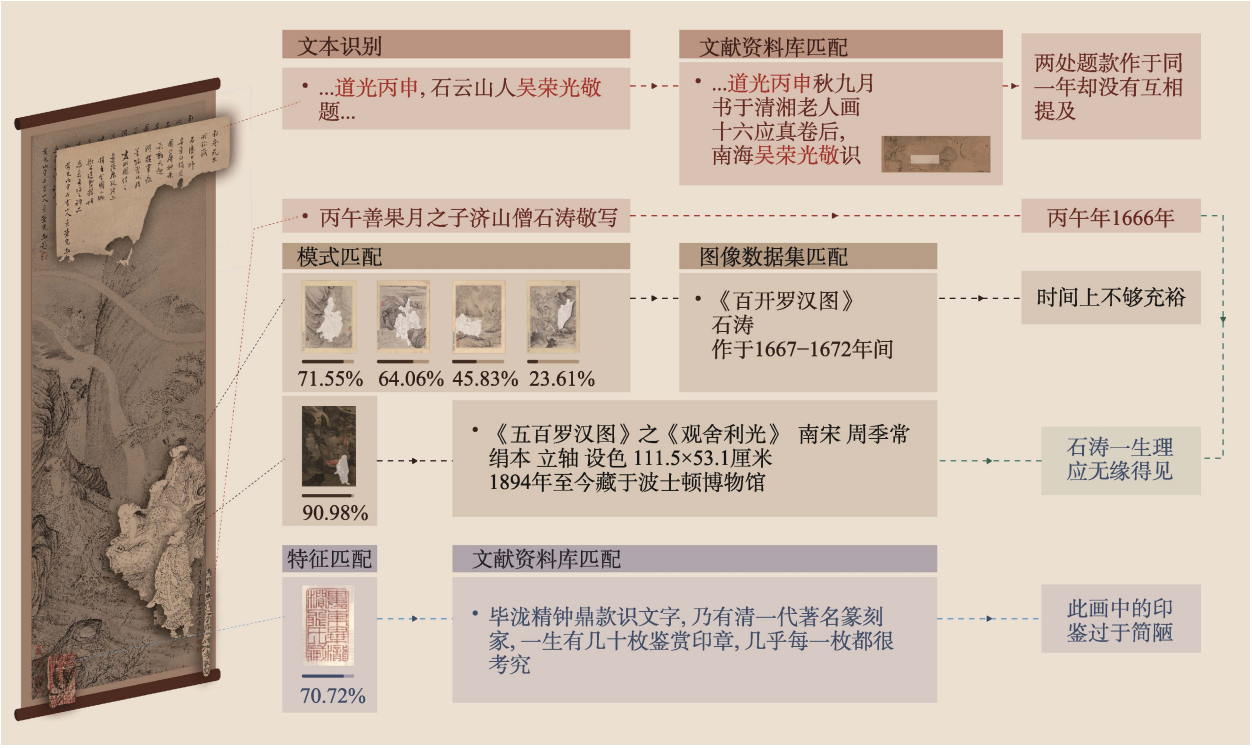


图 8　案例研究关键信息分析

别速度是更快的吗?” “您认为计算机辅助系统对提示的印章考证关联信息数量是更多的吗?”

实验结果如图 9 所示，其中，图 9a 说明计算机辅助鉴定系统对画作的题款、印章、内容都有较强的关联能力，虽然后续验证中存在个别无效关联的条目，但与专家用户相比，其在速度和数量上都有明显优势. 图 9b 展示了专家用户的主观评分结果，色块代表了所选中评分等级，图中通过色块的面积代表了所选中评分的人数情况. 从图 9 可以看出，专家用户整体对计算机辅助系统的认可度比较高，同意及强烈同意的人数占比较大，其中对题款、印章及内容的识别速度和准确率都保持较高的评分，尤其是印鉴是识别和比对是最没有歧义的，加权平均分为 1.78，略有不足的是画作内容的考证关联因为细节复杂且关联数据多影响了效率，使得专家在该方面协作上的最终反馈分数为 0.78.



a. 关联考证条目比较　　b. 专家用户主观评分结果

图 9　用户研究实验结果

5.3　讨　论

在案例研究与专家评估过程中，待鉴定古画图像通过模型的分割、匹配和关系网络构建，再经过专家对于信息的过滤与推断，最终对一幅画的真伪结果形成初步判断. 信息准确的关联十分关键，并且每一处要点都需要文献的辅证，这是从古

画图像本身所不可能获取的. 本系统具备图像分割处理效率高和考证关联信息维度广 2 个特点, 友好的人机协同设计也给专家带来较好的用户体验.

然而, 该系统仍存在一些不足. 首先, 一些关键信息难以通过现有设计的算法进行关联, 这给鉴定带来了一定的挑战. 以本案例为例, 系统所有关联的出发点必然依据题款、印鉴和图像内容三者之一, 但是模型并不能做到语义的理解. 而从专家视角能够直观地联想出这是石涛在罗汉主题的画作, 在石涛的相关生平记录中分别在 1666 年和 1669 年有过相关出游罗汉寺的经历, 专家也许能够从此类相关记录中获取更多信息, 但是这些深层次的关联是本系统无法捕获的.

其次, 大量的画家留存的作品仅仅只有几件或是一件, 真实性完完全全只能通过少数史料来考证. 这意味着需要继续进行系统的优化和升级, 以提高系统在少样本案例中的适用性.

最后, 在计算机辅助鉴定的发展问题上, 本文也收集到了专家的一些反馈. 专家表示, 迄今为止的计算机算法因为存在局限并且缺少沟通, 因此, 双方也少有真正的合作鉴定机会. 但是, 专家认为未来的工作通过计算机科学家和艺术品鉴定专家之间密切的合作和交流, 必然能够取得新的突破.

6 结 语

本文设计了一个基于多模态语义的中国古画可视辅助鉴定的系统. 该系统通过多种技术手段, 实现了对于中国古画画作对象的快速提取, 画作样本匹配和外延考证的快速检索关联分析, 实现了一定程度的自动化辅助鉴定; 并且通过用户可用性测试, 该系统具有较高的鉴定准确性和实用性, 具备较好的人机协同设计功能, 能够为鉴定专业人士提供一定的便利和支持.

未来可以探索更加高效的算法和技术来真正将计算机技术融入鉴定过程中. 随着人工智能相关领域的发展, 多模态模型将会继续突破并更加通用化, 模型将不仅能够分析图像特征, 也能够读懂画作内涵, 甚至通过语音或文本等更简易的交互手段来进行鉴定; 系统在自动化分析图像的同时, 还支持实时在互联网中检索相关资讯, 实现更加智能的鉴定和推理. 这些优化措施将进一步提高中国古画鉴定系统的性能和可靠性, 为相关领域的研究和实践提供更好的支持和帮助.

参考文献(References):

[1] Manovich L. Data science and digital art history[J]. International Journal for Digital Art History, 2015(1): 14-34

[2] Tifentale A, Manovich L. Selfiecity: exploring photography and self-fashioning in social media[M] //Berry D M, Dieter M. Postdigital Aesthetics: Art, Computation and Design. London: Palgrave Macmillan, 2015: 109-122

[3] Guan Xiaohui. Research on computer-aided classification, authentication and system of Chinese traditional painting[D]. Hangzhou: Zhejiang University, 2005(in Chinese)
(关晓惠. 计算机辅助的国画: 分类、鉴别与系统[D]. 杭州: 浙江大学, 2005)

[4] Jia Chunhua, Guo Xiaoying, Bai Ruyi. Review of feature extraction methods and research on affective analysis for paintings[J]. Journal of Image and Graphics, 2018, 23(7): 937-952 (in Chinese)
(贾春花, 郭小英, 白茹意. 绘画特征提取方法与情感分析研究综述[J]. 中国图象图形学报, 2018, 23(7): 937-952)

[5] Fiorucci M, Khoroshiltseva M, Pontil M, *et al*. Machine learning for cultural heritage: a survey[J]. Pattern Recognition Letters, 2020, 133: 102-108

[6] Xu Bangda. Xu Bangda collection 1·Introduction to ancient calligraphy and painting appraisal[M]. Beijing: Forbidden City Publishing House, 2005: 11-13(in Chinese)
(徐邦达. 徐邦达集一·古书画鉴定概论[M]. 北京: 紫禁城出版社, 2005: 11-13)

[7] Chen Xuliang. Authenticity seen in subtle places——the authentication of Huang Binhong's landscapes[J]. Hunan Provincial Museum, 2014(1): 424-440(in Chinese)
(陈叙良. 于精微处见真性——黄宾虹款山水画真伪鉴别[J]. 湖南省博物馆馆刊, 2014(1): 424-440)

[8] Li Daxiang, Zhang Yue. Chinese painting image classification algorithm based on multiscale CNN features[J]. Journal of Xi'an University of Posts and Telecommunications, 2021, 26(1): 104-110(in Chinese)
(李大湘, 张玥. 基于多尺度 CNN 特征的国画图像分类算法[J]. 西安邮电大学学报, 2021, 26(1): 104-110)

[9] Sheng Jiachuan, Li Yuzhi. Learning artistic objects for improved classification of Chinese paintings[J]. Journal of Image and Graphics, 2018, 23(8): 1193-1206(in Chinese)
(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)

[10] Li Yuzhi, Sheng Jiachuan, Hua Bin. Improved embedded learning for classification of Chinese paintings[J]. Journal of Computer-Aided Design & Computer Graphics, 2018, 30(5): 893-900(in Chinese)
(李玉芝, 盛家川, 华斌. 中国画分类的改进嵌入式学习算法[J]. 计算机辅助设计与图形学学报, 2018, 30(5): 893-900)

[11] Shen X, Efros A A, Aubry M. Discovering visual patterns in art collections with spatially-consistent feature learning[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2019: 9270-9279

[12] Pflüger H, Thom D, Schütz A, *et al*. VeCHArt: visually enhanced comparison of historic art using an automated

line-based synchronization technique[J]. IEEE Transactions on Visualization and Computer Graphics, 2020, 26(10): 3063-3076

[13] Cai W. Chinese painting and calligraphy image recognition technology based on pseudo linear directional diffusion equation[J]. Applied Mathematics and Nonlinear Sciences, 2023, 8(1): 1509-1518

[14] Zhu Liangzhi. A study on the authenticity of the extant “Shitao” works[M]. Beijing: Peking University Press, 2017: 65-83(in Chinese)
(朱良志. 传世石涛款作品真伪考[M]. 北京: 北京大学出版社, 2017: 65-83)

[15] Zhu Liangzhi. Research on Shitao[M]. Beijing: Peking University Press, 2005: 509-657(in Chinese)
(朱良志. 石涛研究[M]. 北京: 北京大学出版社, 2005: 509-657)

[16] Khan F S, Beigpour S, van de Weijer J, *et al*. Painting-91: a large scale database for computational painting categorization[J]. Machine Vision and Applications, 2014, 25(6): 1385-1397

[17] Holt B, Weiss K, Niblack W, *et al*. The QBIC project in the department of art and art history at UC Davis[C] //Proceedings of the ASIS Annual Meeting. Waynesville: The Learning & Technology Library, 1997: 189

[18] Bach J R, Fuller C, Gupta A, *et al*. Virage image search engine: an open framework for image management[C] //Proceedings of SPIE 2670, Storage and Retrieval for Still Image and Video Databases IV. Bellingham: Society of Photo-Optical Instrumentation Engineers, 1996, 2670: 76-87

[19] Pflüger H, Ertl T. Sifting through visual arts collections[J]. Computers & Graphics, 2016, 57: 127-138

[20] Gleicher M. Considerations for visualizing comparison[J]. IEEE Transactions on Visualization and Computer Graphics, 2018, 24(1): 413-423

[21] Lawonn K, Trostmann E, Preim B, *et al*. Visualization and extraction of carvings for heritage conservation[J]. IEEE Transactions on Visualization and Computer Graphics, 2017, 23(1): 801-810

[22] Schmidt J, Gröller M E, Bruckner S. VAICo: visual analysis for image comparison[J]. IEEE Transactions on Visualization and Computer Graphics, 2013, 19(12): 2090-2099

[23] Garces E, Agarwala A, Hertzmann A, *et al*. Style-based exploration of illustration datasets[J]. Multimedia Tools and Applications, 2017, 76(11): 13067-13086

[24] Hayem-Ghez A, Ravaud E, Boust C, *et al*. Characterizing pigments with hyperspectral imaging variable false-color composites[J]. Applied Physics A, 2015, 121(3): 939-947

[25] Xu W Z, Tang X J, Zhang G, *et al*. Research on mural painting appreciatione based on spectral imaging and spectral analysis[J]. Spectroscopy and Spectral Analysis, 2017, 37(10): 3235-3241

[26] Lyu S Q, Yang X Y, Pan N, *et al*. Spectral heat aging model to estimate the age of seals on painting and calligraphy[J]. Journal of Cultural Heritage, 2020, 46: 119-130

[27] Swain M J, Ballard D H. Color indexing[J]. International Journal of Computer Vision, 1991, 7(1): 11-32

[28] Stricker M A, Orengo M. Similarity of color images[C] //Proceedings of SPIE 2420, Storage and Retrieval for Image and Video Databases III. Bellingham: Society of Photo-Optical Instrumentation Engineers, 1995, 2420: 381-392

[29] Hsu W, Chua S T, Pung H H. An integrated color-spatial approach to content-based image retrieval[C] //Proceedings of the 3rd ACM International Conference on Multimedia. New York: ACM Press, 1995: 305-313

[30] Haralick R M, Shanmugam K, Dinstein I H. Textural features for image classification[J]. IEEE Transactions on Systems, Man, and Cybernetics, 1973, SMC-3(6): 610-621

[31] Tamura H, Mori S, Yamawaki T. Textural features corresponding to visual perception[J]. IEEE Transactions on Systems, Man, and Cybernetics, 1978, 8(6): 460-473

[32] Lowe D G. Distinctive image features from scale-invariant keypoints[J]. International Journal of Computer Vision, 2004, 60(2): 91-110

[33] Sheng Jiachuan. Automatic categorization of traditional Chinese paintings based on wavelet transform[J]. Computer Science, 2014, 41(2): 317-319(in Chinese)
(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)

[34] Fong W C. Toward a structural analysis of Chinese landscape painting[J]. Art Journal, 1969, 28(4): 388-397

[35] Liu Shang, Sheng Jiachuan. Artists recognition via line shape and ink color distribution of the principal direction for Chinese paintings[J]. Pattern Recognition and Artificial Intelligence, 2017, 30(10): 917-927(in Chinese)
(刘赏, 盛家川. 基于线条形状和主方向墨色分布特征的中国画画家识别算法[J]. 模式识别与人工智能, 2017, 30(10): 917-927)

[36] Ibrahim L M. Design and implementation an automated system for analyzing brushstrokes to distinguish between Van Gogh and his contemporaries by using swarm intelligent method[J]. International Journal of Computer Applications, 2018, 179(11): 5-14

[37] Elgammal A, Kang Y, Den Leeuw M. Picasso, Matisse, or a fake? Automated analysis of drawings at the stroke level for attribution and authentication[C] //Proceedings of the 32nd AAAI Conference on Artificial Intelligence. Palo Alto: AAAI Press, 2018: 42-50

[38] Kirillov A, Mintun E, Ravi N, *et al*. Segment anything[C] // Proceedings of the IEEE/CVF International Conference on Computer Vision. Los Alamitos: IEEE Computer Society Press, 2023: 3992-4003

[39] Qiu Dehong, Chen Chuanbo, Jin Xianji. Signature verification by support vector machine with ALCSL&LS kernel[J]. Journal of Fudan University (Natural Science), 2004, 43(5): 805-809+814(in Chinese)
(邱德红, 陈传波, 金先级. 基于 ALCSL&LS 核函数的支持向量机签名认证(英文)[J]. 复旦学报(自然科学版), 2004, 43(5): 805-809+814)

[40] Yang H M, Lu J J, Lee H J. A Bézier curve-based approach to shape description for Chinese calligraphy characters[C] //Proceedings of the 6th International Conference on Document Analysis and Recognition. Los Alamitos: IEEE Computer Society Press, 2001: 276-280

[41] Lee S, Pan J C. Offline tracing and representation of signa-

tures[J]. IEEE Transactions on Systems, Man, and Cybernetics, 1992, 22(4): 755-771

[42] Yu H Y, Chen J Y, Li B, *et al*. Benchmarking Chinese text recognition: datasets, baselines, and an empirical study[OL]. [2023-06-18]. https://arxiv.org/abs/2112.15093

[43] Chen J Y, Li B, Xue X Y. Scene text telescope: text-focused scene image super-resolution[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2021: 12026-12035

[44] Liang J, Liao W H, Wu Y C. Toward automatic recognition of cursive Chinese calligraphy: an open dataset for cursive Chinese calligraphy text[C] //Proceedings of the 14th International Conference on Ubiquitous Information Management and Communication. Los Alamitos: IEEE Computer Society Press, 2020: 1-5

[45] Zhao Y, Yuan B Z. Skeleton extraction in seal identification system using morphology[C] //Proceedings of the 4th International Conference on Signal Processing. Los Alamitos: IEEE Computer Society Press, 1998: 1281-1283

[46] Sablatnig R, Kammerer P, Zolda E. Hierarchical classification of paintings using face-and brush stroke models[C] //Proceedings of the 14th International Conference on Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 1998: 172-174

[47] Widjaja I, Leow W K, Wu F C. Identifying painters from color profiles of skin patches in painting images[C] //Proceedings of the International Conference on Image Processing. Los Alamitos: IEEE Computer Society Press, 2003: 1-845

[48] Keren D. Painter identification using local features and naive Bayes[C] //Proceedings of the International Conference on Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2002: 474-477

[49] Ufer N, Lang S, Ommer B. Object retrieval and localization in large art collections using deep multi-style feature fusion and iterative voting[C] //Proceedings of the Computer Vision–ECCV Workshops. Heidelberg: Springer, 2020: 159-176

[50] Ufer N, Simon M, Lang S, *et al*. Large-scale interactive retrieval in art collections using multi-style feature aggregation[J]. PLoS One, 2021, 16(11): Article No.e0259718