

基于迁移学习的青花瓷多任务图像识别方法 及其应用研究

王琳达

青岛电影学院美术与动画学院, 山东 青岛 266520

DOI:10.61369/EDTR.2026030015

摘要 : 面向博物馆公共教育与数字化传播, 本文以青花瓷为例, 构建“朝代—器型—纹饰”三维知识框架, 并提出 AI 辅助的多维度识别方案。方法上, 基于迁移学习构建多任务联合预测模型, 同步输出三类标签, 并以 Grad-CAM 等可解释可视化提示转化为观众可操作的观察要点。应用上, 将识别结果嵌入导览提示与对照练习, 形成“识别—讲解—练习”闭环。研究认为, AI 可作为连接专家知识与公众理解的“中介工具”, 但应当明确的是, 其主要作用仍为普及与学习, 而非确认最终鉴定结论, 使用中需遵守数据授权与传播规范。

关键词 : 人工智能; 青花瓷; 博物馆教育; 多任务学习; 数字导览

Research on a Multi-Task Image Recognition Method for Blue-and-White Porcelain Based on Transfer Learning and Its Applications

Wang Linda

School of Art and Animation, Qingdao Film Academy, Qingdao, Shandong 266520

Abstract : Focusing on public education and digital dissemination in museums, this paper takes blue-and-white porcelain as an example to construct a three-dimensional knowledge framework of "dynasty-shape-decoration" and proposes an AI-assisted multi-dimensional recognition scheme. Methodologically, a multi-task joint prediction model is constructed based on transfer learning to synchronously output three types of labels, and interpretable visual cues such as Grad-CAM are translated into observable key points for the audience. In terms of applications, the recognition results are integrated into guided tour prompts and comparative exercises, forming a closed loop of "recognition-explanation-practice". The study suggests that AI can serve as an "intermediary tool" connecting expert knowledge with public understanding. However, it should be clarified that its primary role remains in popularization and learning, rather than confirming final appraisal conclusions, and data authorization and dissemination norms must also be adhered to during its use.

Keywords : artificial intelligence; blue-and-white porcelain; museum education; multi-task learning; digital guided tours

引言

在博物馆展柜前, 观众常被一件青花瓷吸引, 却难以回答: 它大致属哪个朝代? 器型是日用还是陈设? 纹饰又表达了何种题材与寓意? 展示标签往往只给出名称、年代与尺寸, 缺少“如何看懂”的观察路径。文物专家能够快速辨别文物的相关特征, 往往依赖的是长期训练形成的视觉经验与分类经验; 而面向公众普及文物知识的关键, 正是把这些经验转换为可理解、可练习的知识框架。

青花瓷作为中国传统瓷器的主流品种之一, 本身就兼具了鲜明的审美辨识度与文化传播力。原始青花瓷于唐宋已经初见端倪, 其纹饰母题、器型演变与釉彩工艺等, 都蕴含了丰富的文化历史烙印, 青花瓷的特点决定了它特别适合作为公众“视觉识别”训练的切入点。近年来, 随着世界各国对文物数字化与公共文化服务能力建设的大力推进, 博物馆亟需一种既保持学术的严谨性、又易于公众理解的文化传播机制, 使普通大众从“看热闹”逐步走向“懂门道”。

本文主要以“朝代—器型—纹饰”三个维度为核心, 提出一套面向公众教育的人工智能辅助识别系统: 通过建立结构化知识框架, 呈现专业且清晰易懂的知识图谱; 在识别端采用了迁移学习与多任务的联合预测, 形成了对文物的多维度同步解读; 在传播端把识别结果转译为导览提示与对照练习, 构建“识别—讲解—练习”闭环。

项目信息: 青岛电影学院校级科研项目“Heritage AI——基于深度学习技术的中国古代文物多维度智能识别系统与教学应用研究”。

作者简介: 王琳达, 女, 汉族, 讲师, 硕士, 主要研究方向为交互设计、人工智能与数字文博。

一、政策与需求：科技赋能文博公共教育的现实语境

近年来，国家层面持续强调以科技创新支撑文物保护、展示与利用。《“十四五”文物保护和科技创新规划》提出提升文物数字化、展示传播与公众服务能力^[1]；2024年修订的《中华人民共和国文物保护法》进一步为文物合理利用与公共文化服务提供法律保障^[2]。在文旅融合与数字化的传播背景下，数字导览程序、互动装置、微信小程序等成为重要的普及渠道，但目前不少应用仍停留在单一枯燥的信息展示层面，缺少基于视觉特征的信息阐述。

与此同时，故宫博物院数字文物库^[3]、数字敦煌等平台以开放资源、交互展示拓展了公众接触文物、了解文物的渠道^[4]。但对于青花瓷这类尤其需要看细节和做分类的文物，现有的开放资源满足不了大众对眼前瓷器更深层次了解的愿望：观者希望将文物库中的知识转化为即时、可操作的观察提示，构建起文物鉴赏的专业知识脉络。

二、传统参观的局限性与人工智能技术的介入

在传统参观形式下，观众对文物的理解多依赖展览标签与导览人员。展签信息通常高度概括，导览人员的讲解时间与人力分配又难以覆盖每件展品的细节，导致观展者对于眼前文物容易产生“雾里看花”的知识断层。青花瓷的识别涉及到瓷器的廓形气质、花纹构图习惯、口沿与足部的比例、边饰构成的方式等多种知识线索，这些线索本质上可理解为一种能够被大模型训练的视觉能力，但长期以来缺少面向公众的知识谱系的呈现与观察练习的系统。

将人工智能技术引入文物鉴赏的最大意义，不在于用一个绝对标准的答案替代人类对文物的多元化解理解，而在于提供一种新的知识学习途径：它可以将文物专家鉴别经验中可复述的观察要点拆解出来，通过图像识别与可视化的提示词呈现给观众，并与专业知识进行对照校验，从而降低观者对于文物知识学习的门槛、易于进行知识图谱的构建。如图1所示，人工智能赋能模式可在观展者参观时辅助他们对展品进行多维解读。



图1 传统参观模式与 AI 辅助模式下的公众认知路径对比

三、文物知识框架与识别方法：从多维解读到可讲解线索

(一) “朝代—器型—纹饰”三维知识框架

为了使此模型的识别结果能够更加易于讲解与学习，在设计

时围绕青花瓷的特点，将其分类信息概括为“朝代—器型—纹饰”三层维度：朝代维度主要负责提供历史时代信息；器型维度主要负责呈现瓷器的功能与用途；纹饰维度则主要揭示出图案的含义、审美风格与时代工艺特征。三个维度共同构建出公众理解青花瓷知识的简单清晰、更易于理解知识框架，这套知识体系框架既能完美的覆盖展览中常见的问题，也便于在后期研究中扩展到胎釉、款识、窑口等更专业层级的研究。框架示意图2。

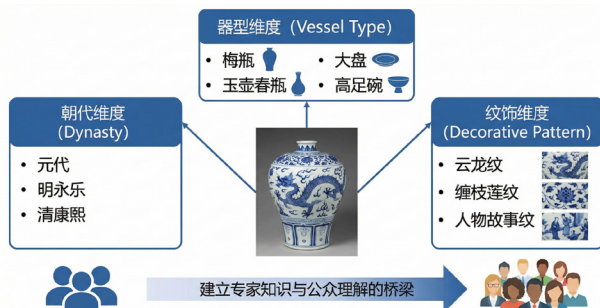


图2 面向公众教育的青花瓷“朝代—器型—纹饰”三维知识结构体系

(二) 迁移学习与多任务联合预测

在识别端，为适应文物图像样本分散、拍摄条件不一的局限性，可采用迁移学习的策略^[5]：以在大规模图像数据集上预训练的卷积神经网络作为特征提取骨干，再针对馆藏与图片进行微调训练^[6]。如图3所示，区别于单一的标签识别，多任务联合预测在一次输入中可同步输出朝代、器型与纹饰三类结果，既明显减少碎片化信息带来的知识理解断裂，也有助于在讲解端形成结构化的知识图谱。

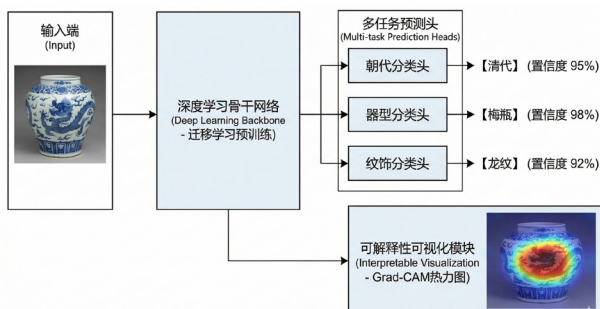


图3 基于迁移学习与多任务联合预测的 AI 识别模型架构

(三) 可解释性提示：把“黑箱结论”变成“观察要点”

作为一项面向社会大众的应用，文物鉴别系统不应仅仅以输出结果为目标。为避免 AI 成为“黑箱结论”，可引入 Grad-CAM 等可解释性可视化方法，将模型判别时聚焦的关键区域以热力图方式呈现^[7]，进而转化为可理解的视觉特点，例如：瓷器口沿与腹壁过渡处的线条、足墙比例、主题纹样的分区方式、边饰的节奏与留白等。需要注意的是，可解释性热力图并不等同于专家鉴定的唯一依据，其价值在于提示观者对于文物鉴赏的视觉切入点，并促进与专业权威知识的对照与验证。

四、场景化转译：构建“识别—讲解—对照练习”闭环

(一) 展厅导览：分层呈现与可视化引导

在展览中，AI 系统可以作为辅助工具，通过嵌入二维码导

览、小程序或互动屏等多个方式与观众互动。参观者在拍摄或锁定关注的展览器物后，系统将分层方式输出信息：第一层给出朝代、器型与纹饰特点的简单判断与关键词提示；第二层进行关键特征的进一步解释；第三层提供可视化的提示，循序渐进的引导观展者对器物进行多维度的系统观察，实现一个普通大众从仅能读懂展牌信息，逐步过渡到在观赏文物中主动去捕捉文物视觉特征，掌握有效信息。

（二）教学实践：题库化练习与反馈机制

在博物馆与学校的合作或课堂教学中，此系统可扩展为“学习—练习—复盘”的任务链形式：首先，围绕经典器型与纹饰母题提供示例与对比图；随后，在练习阶段引导学生先独立判断，再查看系统提示与专业解读；最后的复盘阶段引导学生记录误判的原因、总结可迁移的观察策略。通过此类反复对照练习，系统承担的不再仅仅是给出答案的角色，而是成为一个引导者，促进学生的学习、交流、反馈。图3展示了导览与练习闭环的体验场景。



图4 AI赋能的数字导览与“识别—讲解—对照练习”闭环体验

五、使用边界与规范建议

值得关注的是，文物鉴定涉及器物的材质、工艺、年代、修复痕迹、流转记录等多重线索，单一的图像识别无法覆盖全部线

索。因此，本文展示的系统目前定位在便于历史文化传播、文物知识普及与大众初识别文物的辅助工具，而非文物鉴别专家或专业检测机构工作的替代品。在应用时，我们应避免将人工智能的输出结果作为唯一的权威结论，系统可采用“提示性信息+参考范围+可再确认入口”的呈现策略，并应在界面中明确免责声明。

在数据与治理层面应注意：其一，采集与使用馆藏图像需遵守授权、版权与隐私规范，建立可追溯的数据来源说明；其二，标签体系应保持术语边界清晰，区分“公众可用标签”与“专业细分标签”，避免误导大众；其三，在多个博物馆协作与资源共建中推动元数据标准化，提高数字资源的可复用性与可检索性；其四，针对青少年学习场景，应倡导将AI提示与策展、教育活动相结合，避免学生采取死记硬背式的学习方式。

六、结语

人工智能进入文博领域的意义不在于创造更精准的结果，而在于提供一种知识转译与传播途径：把此领域专家们分散化的知识线索串联为一个可供学习、可供检索、易于解释与理解的知识单元。以青花瓷为例，“朝代—器型—纹饰”的学习框架与多任务识别路径能够为公众提供更为清晰的观察路线，促使观展者从“看见”走向“理解”。此领域未来仍需在数据规范、跨馆协作与教育设计上持续改进与完善，让技术真正服务于文化理解与文化传播。

参考文献

- [1] 国务院办公厅. 国务院办公厅关于印发“十四五”文物保护和科技创新规划的通知[EB/OL]. (2021-11-08) [2026-01-12]. https://www.gov.cn/zhengce/zhengce-ku/2021-11/08/content_5649764.htm.
- [2] 全国人民代表大会常务委员会. 中华人民共和国文物保护法（2024年修订）[EB/OL]. (2024-11-09) [2026-01-12]. https://www.ncha.gov.cn/art/2024/11/9/art_2301_42898.html.
- [3] 故宫博物院. 故宫博物院数字文物库[EB/OL]. [2026-01-12]. <https://digicol.dpm.org.cn/>.
- [4] 敦煌研究院. 数字敦煌[EB/OL]. [2026-01-12]. <https://www.e-dunhuang.com/>.
- [5] LeCun Y, Bengio Y, Hinton G. Deep learning[J]. Nature, 2015, 521(7553): 436-444.
- [6] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016: 770-778.
- [7] Selvaraju R R, Cogswell M, Das A, et al. Grad-CAM: Visual explanations from deep networks via gradient-based localization[C]//Proceedings of the IEEE International Conference on Computer Vision. Venice: IEEE, 2017: 618-626.