

基于深度学习的硬笔书法学习系统设计

刘宏烽¹ 张路遥²

LIU Hongfeng ZHANG Luyao

摘要

针对硬笔书法教学的数字化与智能化需求,文章设计并实现了一款硬笔书法智能学习系统。该系统采用 CNN 与 Vision Transformer (ViT) 双分支网络架构,分别聚焦书写笔画细节与整体结构特征,结合对比学习机制进一步提升评测的精准度;同时集成课程学习、书写训练、闯关挑战、擂台竞技及可视化反馈等功能模块,通过游戏化设计增强用户学习的趣味性与主动性,为硬笔书法教学的数字化、智能化与游戏化融合提供了可行的综合解决方案。

关键词

深度学习;硬笔书法;学习系统;智能评测;游戏化学习

doi: 10.3969/j.issn.1672-9528.2026.01.009

0 引言

近年来,国家越来越重视中小学书法教育,陆续出台了一系列政策文件。2011年,教育部发布了《关于中小学开展书法教育的意见》,首次明确将书法教育纳入中小学教学体系,强调通过书法课程传承中华优秀传统文化。2013年,教育部进一步细化要求,印发《中小学书法教育指导纲要》,针对不同学段制定了具体的书法教学目标和内容标准。到了2024年,教育部出台了《关于进一步加强中小学规范汉字书写教育的通知》,从书写习惯培养、师资建设到评价体系完善等方面提出系统性要求。足见国家对书法教育的重视和支持。然而,书法教育发展受阻于教材和师资两方面的问题,^[1]学生难以提供及时、客观、个性化的指导。随着人工智能技术不断革新进步,与书法教学进一步融合创新,书法教学的数字化转型迎来前所未有的发展契机^[2]。基于此背景,本研究旨在开发一个功能全面、评测精准、体验有趣的智能硬笔书法学习系统。在核心技术上,利用最前沿的 CNN 与 ViT 融合模型^[3-5],同时攻克局部细节与全局结构两大评测难点,实现了客观、精准的自动化评分;在功能设计上,通过闯关、擂台比武等游戏式闯关的方式^[6],激发学习者的内在动机,变“被动练习”为“主动探索”;在学习反馈上,提供实时、可视化的测评结果和指导意见^[7],让反馈清晰易懂。整个系统将先进算法与游戏化学习理念相结合,创造了沉浸式的书法学习体验。

1 系统的总体设计

1.1 系统需求分析

系统的主要用户是小学阶段以及学前幼小衔接段的少儿,该群体处于书写习惯养成的关键期,需要有趣味性的引导和即时正向的反馈来维持学习动力。其对系统的核心需求是基础笔画的规范训练、字形结构的初步掌握以及与语文课本生字同步的练习内容。基于上述用户需求分析以及学情分析,把系统功能设计为课程学习、书写训练、闯关训练、擂台比武、学习反馈和系统设置六大模块。

1.2 系统功能模块设计

系统功能分为课程学习、书写训练、闯关训练、擂台比武、学习反馈和系统设置六大模块。

1.2.1 课程学习模块

内容涵盖了坐姿握笔、幼小衔接、一至六年级每学期语文教材生字的训练课程,以及书法训练的方法和技巧、卷面提分训练、书法创作课程。所有课程以图文、动画和短视频相结合的形式呈现,录制教学视频和示范视频,对重点难点笔画和汉字进行慢动作回放、多角度讲解,具体内容如图1所示。学生完成一个视频的观看,则可以获得“小红花”奖励。

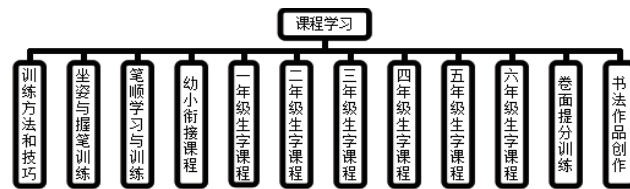


图1 课程学习模块内容

1.2.2 书写训练模块

书写模块提供一至六年级每学期语文教材生字的训练,以及卷面提分、书法作品创作训练,书写训练内容与课程学

1. 广州城建职业学院信息工程学院 广东广州 510925

2. 广州宝露软件开发有限公司 广东广州 510660

[基金项目] 2023年广东省教育厅普通高校重点领域专项课题“基于深度学习的少儿硬笔书法训练纠正系统研发”(2023ZDZX1096)

习模块的内容相对应。学生完成书写后拍照上传到系统，系统生成智能评测报告，反馈分数，指出相应的问题，提出修改意见，为孩子改进书写提供指导。结果为“通过”则可以选择下一个字进行训练或者退出，结果为“不通过”则可以选择重练或者观看教学视频，学生可以选择重新书写，也可以通过观看视频，学习书写技巧之后再重新书写。完成一个字的训练可以得到“小红花”奖励，通过与不通过获得不同数量的“小红花”。书写训练的流程如图2所示。



图2 书写训练流程

1.2.3 闯关训练模块

将书法学习设计成闯关游戏，可以提升书法学习的趣味性，激发孩子学习兴趣。

在关卡的设计上，闯关内容包括笔画关、部首关、结构关三大部分。笔画关由6个基本笔画关和25复合笔画关组成，共31个大关，每个笔画包含若干个例字，每个例字为一小关；部首关由99个大关组成，每个部首包含若干个例字，每个例字为一小关；结构关由32个常大卡组成，每个结构规律包含若干个例字，每个例字为一小关。关卡的设置如图3所示。闯关者须按关卡中的顺序，通过大关下面的所有小关才算通过了一大关，然后才能进入到下一个大关。

在闯关的过程中，学生打开关卡中的例字并进行书写，完成书写之后拍照上传进行测评，如果测评结果是“通过”则可以进入到下一关，如果测评结果为“不通过”，则要重新学习和书写，直到通过为止。每通过一个小关或者一个大关则可以得到相应的数量的“小红花”奖励。

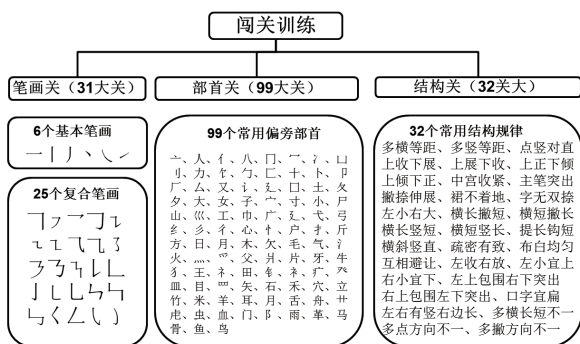


图3 关卡内容设计

1.2.4 擂台比武模块

系统设计了“一对一在线擂台赛”功能，旨在通过竞技形式激发学生的书写兴趣与积极性。系统将根据学生所在年级，智能匹配书写水平相近的在线用户，随机组成一对一竞赛组。书写内容是系统根据学生年龄和书写水平进行推送，确保书写内容与学生水平相匹配。比赛过程中，双方在规定时间内完成

书写并提交作品。系统基于智能评测算法，并生成客观评分。评分较高者获胜，每获胜一局即可获得“小红花”奖励。

此外，系统设立月度“擂主”评选机制。每月统计所有用户在擂台比武中累计获得的“小红花”数量，排名前100的学生将被授予“月度擂主”称号，并获得参与年度总决赛的资格。年度总决赛在每年年底举行，所有月度擂主可同台竞技。总决赛不仅设置更具吸引力的奖品，还将通过线上直播形式增强仪式感与参与感。所有获奖选手除获得丰厚的实物奖励外，还将收到由系统寄出的定制荣誉证书及礼品。通过寓教于乐的方式，有效提升学生的参与感、成就感和书写积极性。

1.2.5 学习反馈模块

学习反馈主要是对学生的进行学习记录，第一是历史学习数据反馈，并进行可视化展示，以图表形式展示生成日度、周度、月度练习报告，汇总练习时长、练习字数、完成关卡数、准确率等数据；第二是学习提升和进步的数据反馈，利用图表形式进行直观化展示，通过对比不同时间段的学习数据，生成个性化进步曲线，可以直观展示书写技能的提升轨迹；第三是获取奖励的数据反馈，包括通过日常练习和擂台比武所获得的“小红花”数量、累计擂主称号次数等。这些成果以“勋章墙”的形式呈现，增强学生的学习成就感与持续练习的动力。

1.2.6 系统设置模块

系统设置模块主要包括个人信息管理、偏好设置和通用系统设置三部分。在个人信息管理方面，用户可查看和修改个人资料，包括头像、昵称、所属学校、年龄、年级等基本信息，以便系统更精准地匹配学习内容与评测标准，同时完善个人学习档案，为学习数据分析提供基础。此外，该模块还提供书写偏好、通知设置、隐私与账号管理等常用系统配置功能，使用户能够根据自身需求定制使用体验。

1.3 系统架构设计

系统设计采用前后端完全分离的分布式架构，可以有效降低了系统各模块间的耦合度，保障系统的可用性、可扩展性以及良好的可维护性。系统在逻辑上划分为四个层次：前端表现层、后端应用层、算法微服务层和数据层，各层组件通过定义良好的标准化接口gRPC进行通信，如图4所示。

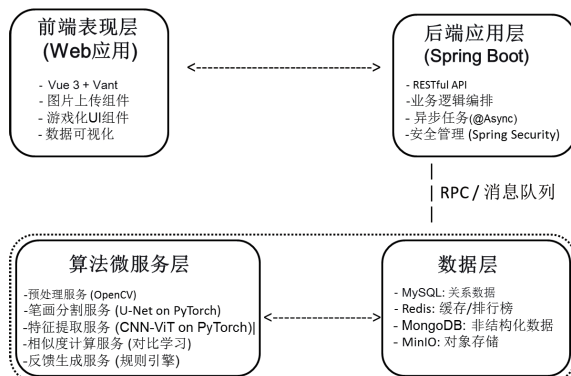


图4 系统架构设计

2 系统功能实现

2.1 前端表现层

系统设计采用 Vue3 组合式 API 作为核心前端框架，构建了一个单页面应用（SPA），利用 Vue Router 实现前端路由管理，Vuex 进行状态管理。UI 组件库选用为移动端打造的 Vue3 组件库 Vant 4，以便获得良好的视觉与操作体验。该层核心功能及实现方式如下：

（1）图片上传组件。基于 Vant Uploader 组件进行开发，集成 EXIF.js 库用于自动校正移动设备拍摄照片的方向问题。使用 Canvas API 实现客户端图片压缩与尺寸调整功能，通过 canvas.toBlob() 方法将图片压缩为指定质量的 JPEG 格式，可以减少网络传输开销。

（2）游戏化 UI 实现。使用 CSS3 动画和 SVG 技术实现勋章点亮、进度条填充等动效。擂台排行榜组件采用 Virtual Scroll 技术优化长列表渲染性能，使用 vue-virtual-scroller 库实现高效渲染。

（3）数据可视化。集成 Apache ECharts 5 库，通过调用其提供的 JavaScript API，将用户的学习数据渲染成交互式图表。使用 echarts.init() 初始化图表实例，并通过 setOption() 方法动态更新图表数据。

（4）状态管理。采用 Pinia 作为 Vue 的官方状态管理库，集中管理用户登录状态、全局配置及跨组件的复杂业务状态。通过 defineStore() 定义 store，使用 ref()、computed() 等响应式 API 管理状态。

前端通过 Axios 库发起 HTTP 请求，与后端 RESTful API 进行数据交互。配置了请求拦截器自动添加 JWT 令牌，响应拦截器统一处理错误信息。

2.2 后端应用层

后端应用层是系统的业务中枢与网关，负责处理所有核心的业务逻辑、流程编排和事务管理。本层采用 Java Spring Boot 3 框架进行构建，利用其自动配置、starter 依赖等特性快速搭建生产级应用。该层的关键实现细节如下：

（1）RESTful API 设计与实现。使用 Spring Web MVC 模块构建 REST 控制器（@RestController）。采用统一 API 响应体封装（Result<T>）和全局异常处理器（@ControllerAdvice + @ExceptionHandler）来规范所有接口的输出格式和错误处理。使用 Swagger UI 通过 @Operation、@Parameter 等注解自动生成 API 文档。

（2）业务逻辑编排（Orchestration）：应用层充当了“流程指挥官”的角色。以提交评测请求为例，其编排流程为：第一步，接收前端上传的图片，通过 MinIO Java SDK 的 putObject() 方法暂存至对象存储；第二步，通过 gRPC Java 客户端同步调用算法层的预处理服务，获取处理后的图像；第三步，使用 Spring 的 @Async 注解异步调用笔画分割与特征提取服务；第四步，使用 CompletableFuture 组合异步结果，

生成最终评测报告；第五步，通过 Spring Data JPA 持久化结果至 MySQL。

（3）异步化与解耦。配置自定义线程池执行 @Async 任务，通过 @EnableAsync 启用异步支持。对于非实时性任务，使用 RabbitTemplate 向 RabbitMQ 队列发送消息，实现系统解耦。

（4）安全与认证：集成 Spring Security 6 配置安全过滤器链，使用 JWT 库实现 JWT 令牌的生成与验证。通过 JwtAuthenticationFilter 实现令牌认证，@PreAuthorize 注解实现方法级权限控制。

2.3 算法微服务层

算法微服务层是系统的智能核心，所有复杂的机器学习模型推理均在此完成。该层使用 Python 3.10+ 和 FastAPI 框架构建，利用 FastAPI 的高性能和自动 OpenAPI 文档生成特性。各微服务的具体技术实现如下：

（1）预处理服务。基于 OpenCV 和 Pillow 库实现。服务使用 cv2.imdecode() 读取图像，依次执行：颜色空间转换（cv2.cvtColor(img, cv2.COLOR_BGR2GRAY)）、高斯模糊去噪（cv2.GaussianBlur()）、大津法二值化（cv2.threshold()）、轮廓查找与最小外接矩形矫正（cv2.findContours(), cv2.minAreaRect()），最终输出归一化后的标准图像。

（2）笔画分割服务。使用 PyTorch 加载预训练好的 U-Net 语义分割模型^[8]。服务使用 ONNX Runtime 作为推理引擎，通过 onnxruntime.InferenceSession() 加载 ONNX 模型，使用 run() 方法执行推理，输出每个笔画的二值掩膜。

（3）特征提取与相似度计算服务：同时加载 CNN^[3] 和 ViT^[5] 两个预训练模型。使用 PyTorch Grad-CAM 生成可视化热力图。

（4）服务通信与部署：使用 gRPC Python 实现服务间通信，通过 Protocol Buffers 定义服务接口。每个服务使用 Docker 容器化，通过 Kubernetes Deployment 进行部署，配置 Horizontal Pod Autoscaler 实现自动扩缩容^[9]。

2.4 数据持久层

数据持久层采用多范式数据库混合方案，针对不同数据的特性和访问模式选用最优的存储引擎。

（1）MySQL 8.0：作为核心业务的关系型数据库，用于存储所有需要事务保证和复杂查询的结构化数据。包括：用户账户信息、系统权限配置、课程与字帖的元数据、关卡设计配置、擂台比赛的元数据（如参与记录、名次）等。使用 Spring Data JPA 作为 ORM 框架，通过 @Entity 定义实体类，@Repository 定义数据访问接口。配置 HikariCP 连接池优化数据库连接管理。建立复合索引优化查询性能。

（2）Redis 7.0：作为高性能的内存键值数据库，扮演着两个关键角色：一是作为缓存层，缓存频繁访问的热点数据（如用户信息、关卡信息），极大地减轻 MySQL 的负

载,提升接口响应速度;二是用作实时排行榜的存储引擎,利用其内置的 ZSet (有序集合) 数据结构,能够高效地处理和更新擂台比武中的用户排名。使用 Spring Data Redis 的 RedisTemplate 进行操作。使用 opsForZSet().add() 方法维护排行榜,opsForValue().set() 缓存热点数据,配置合理的过期时间(TTL)。

(3) MongoDB 6.0: 作为文档型数据库,用于存储结构灵活、schema 易变的非结构化或半结构化数据。使用 Spring Data MongoDB 的 MongoTemplate 进行操作。通过 @Document 注解定义文档实体,使用 MongoRepository 实现复杂查询,为常用查询字段建立索引。

(4) MinIO: 是一个高性能、开源、与 Amazon S3 兼容的对象存储服务。用于安全可靠地存储系统中的所有的大型二进制文件,包括:用户上传的原始书写图片、系统内置的标准字帖图片、算法模型文件(如 .pt 或 .h5 文件),以及服务生成的热力图和报告 PDF 等。使用 Java SDK 的 MinioClient.putObject() 存储用户上传的图片,生成预签名 URL (presignedGetObject()) 供前端访问,配置生命周期策略自动清理临时文件。

3 关键算法模块详细设计

3.1 数据预处理

预处理是保障算法准确性的第一步。该服务接收原始图像,首先,使用 OpenCV 将图像转为灰度图,并采用大津法(Otsu's Method)进行自适应阈值二值化,分离笔迹与背景;然后,使用形态学操作(如开运算)去除小噪点;最后进行尺寸归一化处理,该服务被封装为独立的 Python 服务,通过 FastAPI 等框架提供 API 接口。

3.2 CNN-ViT 双分支特征融合网络

该模型是系统实现精准评测的核心。首先,采用一个轻量化 Custom CNN 作为主干。输入预处理后的图像,通过一系列卷积和池化层,输出一个描述局部细节的特征向量 $\mathbf{F}_{\text{cnn}}$ 。然后,将图像分割成 16×16 的 patches,线性投影后加入位置编码,输入一个标准 ViT-Base 模型的编码器。取 [CLS] token 的输出作为全局特征向量 $\mathbf{F}_{\text{vit}}$ 。最后,采用加权拼接方式将上述两个特征进行融合,用公式表示为:

$$\mathbf{F}_{\text{fusion}} = \text{FC}([\alpha \cdot \mathbf{F}_{\text{cnn}}; \beta \cdot \mathbf{F}_{\text{vit}}]) \quad (1)$$

式中: $\alpha \cdot \mathbf{F}_{\text{cnn}}$ 和 $\beta \cdot \mathbf{F}_{\text{vit}}$ 为可学习的缩放参数; $[\cdot]$ 为拼接操作; FC 为一个全连接层,用于融合和降维。该融合策略能自适应地平衡局部与全局特征的贡献。

3.3 基于对比学习的相似度度量

为了训练网络能够准确评估书法字的质量,采用对比学习的深度学习范式^[10]。该方法的核心在于让模型通过比较学习,理解样本间的相似与差异关系。采用三元组采样(Triple

Sampling) 技术构建训练样本。具体而言,为每个训练样本构建一个三元组 (A, P, N) , 包括 Anchor (A): 标准数字图像,作为参考基准; Positive (P): 与 A 为同一汉字且书写质量优良的样本; Negative (N): 与 A 为同一汉字,但书写质量较差,或完全不同汉字的样本。使用三元组损失(Triplet Loss) 作为训练目标函数,其数学表达式为:

$$L_{\text{circle}} = \log \left[1 + \sum_{j=1}^L \exp(\gamma \alpha_n^j (s_n^j - \Delta_n)) \sum_{i=1}^K \exp(-\gamma \alpha_p^i (s_p^i - \Delta_p)) \right] \quad (2)$$

这种基于对比学习的相似度度量方法,通过让模型在训练过程中直接学习区分优质和劣质书写样本,使其获得了接近人类专家的书法审美判断能力,实现了客观、准确的自动化书法评价。

3.4 U-Net 笔画分割与笔画级评测

采用标准的 U-Net 架构,其编码器-解码器结构结合跳跃连接,能精准定位每个笔画像素。输出为与输入图像同尺寸的多通道分割图,每个通道对应一个笔画类别。利用分割结果掩膜,从原图中裁剪出每一个笔画。将用户的一个笔画图像与范字的对应笔画图像,送入此前已训练好的双分支网络和相似度计算模块,得到该笔画的独立评分。这意味着整体和笔画评测共享同一个强大的特征提取器,保证了评价标准的一致性。

3.5 可视化反馈与意见生成

3.5.1 可视化反馈

使用 Grad-CAM++ 技术。对双分支网络最后的卷积层特征图进行加权求和,生成针对“整体字”的类激活热力图,直观显示哪些区域导致了低分,实现可视化反馈。

3.5.2 文本意见生成

构建一个规则库。规则由书法专家知识定义,与模型输出绑定。

IF “整体结构评分” < 阈值 AND “ViT 分支的注意力权重”, 显示左重右轻, THEN 提示: “整体结构左紧右松, 请注意右边部分的笔画应向外展开一些。”

IF “撇画评分” < 阈值, THEN 提示: “该撇画书写过于僵硬, 缺乏流畅的弧度。”系统根据评测结果触发相应规则, 填充模板, 生成自然语言反馈。

4 结语

本文设计了一个集智能评测算法与游戏化设计于一体的硬笔书法学习系统。系统通过 CNN-ViT 双分支融合网络 and 对比学习机制,实现了从笔画到结构的多维度分析与测评。通过构建“课程-训练-闯关-擂台-反馈”的功能闭环,创造了有趣、沉浸式的学习体验,有效激发了学习者的内在动机。为大规模、低成本的书法智能化教学推广提供了坚实的理论和实践基础。

SCFSA-YOLO：遮挡下精准客流统计的实时头肩检测器

王浩¹

WANG Hao

摘要

密集场景下的客流统计面临小目标信息量匮乏、目标遮挡严重的现实挑战，导致现有检测方法难以满足实际应用需求。为此，文章提出一种改进的头肩检测模型 SCFSA-YOLO，该模型以 YOLOv12 为基础框架，通过针对性优化小目标检测性能，有效提升复杂密集环境下客流统计的准确性。首先，选用头肩作为检测目标以缓解遮挡问题，并在 VSCrowd 数据集上完成模型训练；其次，通过增设小目标专用检测层、加深主干网络结构，以及将输入分辨率从 640 提升至 960，显著增强小目标的特征表征能力；此外，设计空间多尺度 - 通道频率协同注意力机制（SCFSA），提升模型在遮挡环境下的空间定位精度；同时，提出增强型空间 - 光谱下采样模块（ESSD），替代传统步长卷积，确保下采样过程中原始特征信息的最大化保留。实验结果表明，所提方法在 VSCrowd 测试集上的 mAP@0.5 达到 90.0%，较基准 YOLOv12 模型提升 8.6%，在小目标检测和遮挡场景中展现出明显优势。综上所述，SCFSA-YOLO 模型通过架构优化与创新的注意力机制、下采样，有效解决了密集场景中小目标表征不足和严重遮挡的核心难题，为高精度、实时化的客流统计应用提供了切实可行的技术方案。

关键词

头肩检测；密集场景；YOLOv12；小目标检测；注意力机制；客流统计

doi: 10.3969/j.issn.1672-9528.2026.01.010

0 引言

随着社会的发展，出行旅游、购物越来越普遍，公共场所如车站、商场、景区等人员聚集现象增多，密集人群带来

的安全隐患不容忽视。据统计，近年来全球范围内因人群拥挤导致的踩踏事故屡见不鲜，造成大量人员伤亡。因此，实时精准的客流统计与密度分析对于公共安全管理 and 突发事件预警具有至关重要的意义。

传统的客流统计方法主要依赖人工观察或简单的传感器

1. 北京数原数字化城市研究中心 北京 100083

参考文献：

- [1] 向彬, 陈昂. 我国中小学阶段书法教育的现状 [J]. 中国艺术, 2017(3): 16-21.
- [2] 刘宏烽, 张路遥. 人工智能背景下书法教学数字化转型：动因、挑战与进路 [J]. 宿州职业技术学院学报, 2025(4): 56-62.
- [3] 温佩芝, 姚航, 沈嘉伟. 基于卷积神经网络的石刻书法字识别方法 [J]. 计算机工程与设计, 2018, 39(3): 867-872.
- [4] 郭佳霖, 智敏, 殷雁君, 等. 图像处理中 CNN 与视觉 Transformer 混合模型研究综述 [J]. 计算机科学与探索, 2025, 19(1): 30-44.
- [5] 胡正南, 胡立坤. 基于 Vision Transformer 多模型融合的视觉闭环检测算法 [J]. 激光杂志, 2024, 45(6): 75-81.
- [6] 李玉斌, 宋金玉, 姚巧红. 游戏化学习方式对学生学习效果的影响研究：基于 35 项实验和准实验研究的元分析 [J]. 电化教育研究, 2019, 40(11): 56-62.

- [7] 胡智丹, 王志军, 王萌, 等. 基于深度学习的汉字硬笔楷书智能评测系统的设计与应用 [J]. 数字教育, 2025, 11(3): 30-36.
- [8] 肖政文, 刘丹. 基于优化 U-net 的单色图像分割算法 [J]. 信息记录材料, 2025, 26(1): 204-207.
- [9] 汪嘉珮. 基于微服务架构的在线教育系统设计 with 实现 [D]. 武汉：华中科技大学, 2020.
- [10] 刘子扬, 王朝坤, 章衡. 图对比学习方法综述 [J]. 软件学报, 2026(1): 180-199.

【作者简介】

刘宏烽（1981—），男，广东韶关人，硕士，副高，研究方向：信息技术、书法教学。

张路遥（1983—），男，辽宁辽阳人，硕士，工程师，研究方向：软件技术、人工智能。

（收稿日期：2025-11-06 修回日期：2026-01-12）